N° d'ordre : 53

THÈSE

PRÉSENTÉE

A LA FACULTÉ DES SCIENCES DE L'UNIVERSITÉ DE BESANÇON

POUR OBTENIR

LE TITRE DE DOCTEUR EN MATHÉMATIQUES

Mention : MATHÉMATIQUES PURES

PAR

Jean GACHES

Licencié ès Sciences

THÈSE

COMPATIBILITÉ D'UNE SOLUTION CALCULÉE AVEC LES DONNÉES D'UN SYSTÈME LINÉAIRE A COEFFICIENTS INCERTAINS

Soutenue le 13 Septembre 1966, devant la Commission d'Examen

JURY : MM. **CHATELET,** *Professeur* — **Président**

GASTINEL, *Professeur*

RIGAL, *Professeur* — **Examinateurs**

PARIZET, Maître de Conférences

IMP. H. ROBERT – BESANÇON

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41547626_0011

N° d'ordre : 53

THÈSE

PRÉSENTÉE

A LA FACULTÉ DES SCIENCES DE L'UNIVERSITÉ DE BESANÇON

POUR OBTENIR

LE TITRE DE DOCTEUR EN MATHÉMATIQUES

Mention : MATHÉMATIQUES PURES

PAR

Jean GACHES

Licencié ès Sciences

THÈSE

COMPATIBILITÉ D'UNE SOLUTION CALCULÉE AVEC LES DONNÉES D'UN SYSTÈME LINÉAIRE A COEFFICIENTS INCERTAINS

Soutenue le 13 Septembre 1966, devant la Commission d'Examen

JURY : MM. **CHATELET,** *Professeur* — **Président**
GASTINEL, *Professeur*
RIGAL, *Professeur*
PARIZET, Maître de Conférences — **Examinateurs**

IMP. H. ROBERET – BESANÇON

UNIV.
LEIDEN
BIBL.

UNIVERSITE DE BESANCON

FACULTE DES SCIENCES

Doyen Honoraire : M. GLANGEAUD

Professeurs Honoraires MM. BAILLAUD, DUFFIEUX, THIRY, TRILLAT, UEBERSFELD, VINCENSINI

Maître de Conférences Honoraire : M. GRANIER

DOYEN	M.	JACQUEMAIN	Chimie Appliquée
ASSESSEUR	M.	MESNAGE	Chronométrie
PROFESSEURS	MM.	TRONCHET	Botanique
		CHATELET	Mathématiques
		PERROT	Chimie Générale
		DREYFUSS	Géologie Minéralogie
		THEOBALD	Paléontologie et Géologie stratigraph.
		QUANTIN	Botanique
		CHALEAT	Mécanique Appliquée
		NAHON	Mécanique générale
		BERNARD	Chimie- Physique
		RIPPLINGER	Physiologie animale
		RIGAL	Calcul Numérique
		VIENOT	Physique
		THIEBAUT	Géologie Minéralogie
		GOMOT	Zoologie
		PLUVINAGE	Physique théorique
		ARBEY	Astronomie
		MOTCHANE	Electronique
PROFESSEUR ASSOCIE	M.	DEFIX	Mécanique appliquée
PROFESSEURS SANS CHAIRE	MM.	POTIER	Physique des Solides
		DEVIN	Chimie générale
	Mle	LEGRAND	Botanique
	MM.	REAL	Zoologie
		VALLEE	Mathématiques appliq.
		GOUARNE	Mathématiques
		CERUTTI	Mathématiques
MAITRES DE CONFERENCES	MM.	THEOBALD J.G.	Spectroscopie Hertz.
		BONNEAU	Astronomie
		GUERIN	Mathématiques
		PARIZET	Mathématiques
CHARGES D'ENSEIGNEMENT	M.	REMOND	Chronométrie
CONSEILLER ADMINISTRATIF UNIVERSITAIRE	M.	LAMBERT	

à mes PARENTS

pour les sacrifices consentis

à ma FEMME

à mon FILS

Je tiens à assurer de toute ma reconnaissance :

Monsieur CHATELET, Professeur à la Faculté des Sciences de BESANCON qui m'a fait l'honneur de présider le Jury de cette thèse.

Monsieur GASTINEL, Professeur à la Faculté des Sciences de GRENOBLE qui m'a initié à l'Analyse Numérique, me donnant le goût des Mathématiques Appliquées sans jamais ménager ses efforts, et m'a toujours accueilli très chaleureusement dans son laboratoire.

Monsieur RIGAL, Professeur à la Faculté des Sciences de BESANCON qui m'a inspiré l'objet de ce travail et qui, tant par ses critiques que par ses suggestions, en a dirigé le développement et permis la réalisation.

Monsieur PARIZET, Maître de Conférences qui a bien voulu accepter de faire partie du Jury de cette thèse.

Je remercie Monsieur ROBERT, Chercheur C.N.R.S. du Laboratoire de Calcul de Grenoble pour l'aide et les encouragements qu'il m'a apportés au cours de mes recherches.

-=-=-=-=-=-=-=-=-

S O M M A I R E

Le chapitre I est consacré à l'étude des normes de vecteurs et de matrices (applications de C^n et C^{mxn} dans R_+). La caractérisation géométrique en est faite à partir de convexes équilibrés. Une étude analogue et plus générale présente la notion de normes vectorielles de vecteur (application de C^n dans R_+^k).

Le chapitre II expose dans une première partie, quelques lemmes fondamentaux concernant l'ensemble des vecteurs $A\tilde{x}$ où $\tilde{x}$ est un vecteur donné et A une matrice décrivant un lieu convexe (défini par une norme Spq , de Holder ou vectorielle). Nous en déduisons immédiatement des conditions d'existence d'une matrice singulière dans le lieu A^* des matrices "indiscernables" d'une matrice donnée A_o.

La deuxième partie de ce chapitre traite de la validité d'une solution calculée d'un système linéaire dont les éléments ne sont connus qu'avec une certaine incertitude. Le résultat $\tilde{x}$ est accepté comme solution du système (A_o, b_o) s'il existe un système indiscernable (A, b) admettant $\tilde{x}$ solution exacte Les lemmes de la partie précédente permettent de donner des conditions nécessaires et suffisantes simples d'acceptabilité d'un résultat. Une application aux racines d'un polynôme en découle.

Le chapitre III est consacré à des essais numériques effectués sur CAB 500 portant sur la résolution des systèmes linéaires par des méthodes itératives

Le fait de chercher un vecteur x_p, solution exacte d'un système appartenant au lieu d'incertitude apporte un gain sur le nombre d'itérations et ce quelle que soit la méthode utilisée.

I N T R O D U C T I O N

Après avoir appliqué un algorithme -que l'on suppose "résoudre" un type de problème- à un ensemble de données, il importe de connaître la signification du résultat numérique obtenu, c'est-à-dire, de préciser dans quelle mesure ce résultat est acceptable compte tenu du problème proposé.

Cette question est indépendante du temps et du coût d'exécution de l'ensemble des calculs effectués.

Nous ne nous intéressons ici qu'au résultat et non à la manière dont il a été obtenu (bien entendu une étude complète tiendrait compte à la fois du coût et de la signification de la réponse fournie par le critère d'analyse).

La solution exacte, (plus exactement l'ensemble des solutions exactes) ne peut être atteinte que tout à fait exceptionnellement (dans certains problèmes d'analyse combinatoire, de gestion ou de comptabilité portant sur des entiers sur lesquels on effectue un nombre peu élevé d'additions, de multiplications et de comparaisons) tant par suite des erreurs de chute (on remplace un algorithme infini par un algorithme fini) que de celles d'arrondi, (dûes au nombre nécessairement fini et constant de chiffres significatifs acceptés par la machine).

Considérons le système $A_o x = b_o$ et $\tilde{x}$ une solution calculée.

$$\text{Posons} \quad \tilde{r} = A_o \tilde{x} - b_o = A_o (\tilde{x} - x)$$

1. La quantité $q(\tilde{x} - x)$ ne pourra être estimée que d'une manière tout à fait théorique et en général très lâche, à l'aide de la double inégalité suivante

$$\boxed{p(\tilde{r}) \,/\, S_{pq}(A_o) \leq q(\tilde{x} - x) \leq S_{qp}(A_o^{-1})\, p(\tilde{r})}$$

qui provient du fait que :

$$\inf_{\substack{y \in C^n \\ y \neq o}} \frac{p(A_o y)}{q(y)} = \frac{1}{\sup\limits_{\substack{z \in C^n \\ z \neq o}} \frac{q(A_o^{-1} z)}{p(z)}} = [S_{qp}(A_o^{-1})]^{-1} \leq \frac{p|A_o(\tilde{x}-x)|}{q(\tilde{x}-x)} \leq \sup_{\substack{y \in C^n \\ y \neq o}} \frac{p(A_o y)}{q(y)} = S_{pq}(A_o)$$

2. Nous pouvons établir cependant un rapport entre l'erreur relative minimale E_m et maximale E_M commise sur $\tilde{x}$, en considérant

$$\frac{q(\tilde{x}-x)}{q(\tilde{x})} [S_{qp} A_o^{-1}]^{-1} \leq \frac{p[A_o(\tilde{x}-x)]}{q(\tilde{x}-x)} \frac{q(\tilde{x}-x)}{q(\tilde{x})} \leq [S_{pq}(A_o)] \frac{q(\tilde{x}-x)}{q(\tilde{x})}$$

ce qui conduit à :

$$\boxed{\frac{E_m}{E_M} = 1 \,/\, S_{pq}(A_o) \,.\, S_{qp}(A_o^{-1}) = \Gamma(A_o)}$$

Ce rapport est appelé conditionnement de la matrice A_o. Une étude de la notion de conditionnement a déjà été faite par NEUMANN [1] , TURING [2] et plus récemment GASTINEL [3]

Il faut aussi tenir compte de la nécessaire incertitude des données qui ne peuvent "entrer en machine" que sous une convention à précision finie.

Il apparaît donc logique de caractériser la qualité des résultats obtenus

I. L'analyse directe

La comparaison du vecteur $\tilde{x}$ (tout ensemble de résultats, convenablement ordonnés, peut être considéré comme un vecteur $\tilde{x}$, de même que l'on peut parler du vecteur "a" des données, $a \in A$ ensemble des données - machine possibles) avec la solution théorique x, peut être estimée à partir précisément de l'erreur commise $(x - \tilde{x})$. Ce dernier vecteur ne peut être connu exactement -sinon $x=(x-\tilde{x})+\tilde{x}$ le serait- mais on peut en estimer l'ordre de grandeur, généralement sous forme d'une majoration du type

$$|x-\tilde{x}| \leq \vec{\varepsilon}$$ (x, $\tilde{x}$, $\vec{\varepsilon}$ vecteurs; comparaison portant composante à composante).

ou plus généralement

$$N\,|x-\tilde{x}| \leq \varepsilon$$

Ce type d'analyse est classique et porte le nom d'analyse directe. Il permet des réponses du type : "*combien de décimales (d'ubis) sont exactes dans les résultats ?*" mais il est délicat à manier et ne conduit généralement qu'à de grossières majorations.

Nous donnons ci-contre une étude de cette majoration dans le cas d'un système linéaire.

II. L'analyse rétrograde considère, elle, la topologie des données et se pose la question suivante : *"De quel problème, $\tilde{x}$ est-il exactement solution"?*.

Montrons la différence entre ces deux types d'analyse, sur un exemple simple : une opération arithmétique $\circ$, "simulée" par une opération-machine $\bullet$ portant sur des données a et b représentées dans la convention arithmétique utilisée par l'ordinateur. On peut interpréter de deux manières différentes l'opération $a \bullet b$

Analyse directe : $a \bullet b = a \circ b + \varepsilon \qquad |\varepsilon| \leq \eta$ fixé

Il existe une incertitude ε sur le résultat

Analyse rétrograde : $a \bullet b = a' \circ b' \qquad |a'-a| \leq \ell_1(\eta)$

$|b'-b| \leq \ell_2(\eta)$

On remplace une opération "simulée" sur des nombres exacts par l'opération arithmétique correspondante portant sur des données approchées.

Il setrouve que l'analyse rétrograde est en général élégante et relativement facile (WILKINSON), en particulier la solution exacte ne joue aucun rôle et n'a donc pas à être connue ni même estimée.

Notons cependant deux inconvénients :

a) il y a en général une infinité de "pseudo-données" a' et b' (les fonctions ℓ_1 et ℓ_2 étant indéterminées).

b) là encore on ne peut que majorer -en général fortement- les incertitudes et les erreurs.

III. Les techniques analysant le résidu

Soit à résoudre le système linéaire

(1) $A_o x = b_o$

et soit $\tilde{x}$ le résultat-machine

Formons :

$$r(\tilde{x}) = A_o \tilde{x} - b_o \qquad (r=o \Leftrightarrow \tilde{x} \text{ solution exacte})$$

Si r est inférieur à l'incertitude sur b, la solution $\tilde{x}$ peut être considérée comme acceptable dans la mesure où $\tilde{x}$ vérifie le système exact

(2) $A_o \tilde{x} = b_o + r$ (b+r étant acceptable comme second membre)

De telles techniques sont très utilisées, ne serait-ce que parce que le calcul de r et de sa grandeur -définie par une norme N - est très facile (éventuellement, on calcule en double précision les composantes de Ax, chacune étant calculée comme produit scalaire $r_i = \sum_j a_{ij} \tilde{x}_j - b_i$)

Notons cependant quelques limites d'une telle analyse.

a) r n'est connu qu'avec une certaine approximation (Π par exemple, si le produit scalaire susnommé est conduit en double précision avec un seul arrondi final) peu simple, en général, notamment dans le cas où l'incertitude sur b_o est très supérieure à Π. (b_o d'origine physique)

Tel n'est pas le cas lorsque le second membre est un nombre connu exactement et dont la seule incertitude provient de sa traduction par un nombre-machine qui a pour incertitude précisément Π.

b) Cette analyse suppose b_o incertain mais A_o infiniment certain ce qui doit être exclu (puisque A_o est représenté par des nombres-machine) et ce qui est assez arbitraire en ce que A_o et b_o jouent un rôle indifférent.

IV - L'analyse à postériori

Dans la plupart des cas, on part de données incertaines, l'ensemble des problèmes P "indiscernables" du problème proposé P_o, étant défini par

$$P \text{ indiscernable de } P_o \iff F(P, P_o) \leq \varepsilon$$

(P_o données représentables en machine ; F fonction caractérisant "l'écart" entre P et P_o ; ε fixant l'ordre de grandeur de l'incertitude).

Nous obtenons après un traitement plus compliqué, impliquant nécessairement des incertitudes et des erreurs -voire même des fautes- un résultat $\tilde{x}$. Si l'on peut trouver -ou démontrer- qu'il existe un problème P indiscernable de P_o (c'est-à-dire tel que $F(P, P_o) \leq \varepsilon$) et admettant $\tilde{x}$ comme solution exacte, nous accepterons $\tilde{x}$ comme solution et nous dirons que $\tilde{x}$ *est compatible avec les données du problème*.

La différence avec l'analyse rétrograde porte sur le fait que nous cherchons ici un problème indiscernable, et non un problème "quelconque", admettant $\tilde{x}$ comme solution.

Tel est le principe de l'"analyse à postériori",

ET NOTRE THESE VISE A DONNER QUELQUES THEOREMES SIMPLES MONTRANT COMMENT ON PEUT PROCEDER A UNE TELLE ANALYSE POUR DES SYSTEMES LINEAIRES ET CE, DANS DES CAS TRES GENERAUX OU AUCUNE AUTRE ANALYSE NE S'AVERE REALISABLE.

Pour de tels problèmes, les lieux"d'indiscernabilité" seront convexes et définis par :

$$(A,b) \text{ indiscernable de } (A_o,b_o) \Longleftrightarrow \left| \begin{array}{l} N_1 (A-A_o) \leq \alpha \\ N_2 (b-b_o) \leq \beta \end{array} \right.$$

(N_1 et N_2 normes de matrice et de vecteur)

et une solution calculée $\tilde{x}$ sera compatible avec les données du système (A_o, b_o) si il existe un couple (A,b) tel que :

$$\left| \begin{array}{l} A\tilde{x} = b \\ N_1(A-A_o) \leq \alpha \\ N_2(b-b_o) \leq \beta \end{array} \right.$$

Dans cette analyse nous tenons compte non seulement de l'erreur commise sur le vecteur b_o (comme dans l'analyse rétrograde) mais aussi de l'erreur sur A_o.

Nous ne considérerons pas qu'il existe un problème à résoudre mais une infinité de problèmes indiscernables et nous viserons à un test simple d'acceptation ou de rejet d'un résultat $\tilde{x}$, indépendamment de tout mode d'obtention de ce résultat.

Chapitre I

I - NORMES ET SEMI-NORMES SUR C^n

I.1 - Définitions

I.1.1 - Semi-norme

Une application $p : C^n \longrightarrow R_+$ sera une semi-norme sur C^n si elle vérifie les axiomes suivants :

(i) $\forall \lambda \in C$: $p(\lambda x) = |\lambda| \, p(x)$ [en particulier : $p(o) = o$]

(ii) $\forall x, y \in C^n$: $p(x + y) \leqslant p(x) + p(y)$

I.1.2 - Norme

Une semi-norme sera une norme si elle vérifie l'axiome supplémentaire suivant :

(iii) $p(x) = o \Longrightarrow x = o$ [avec i) : $p(x) = o \Longleftrightarrow x = o$]

Correspondant à cette norme, l'écart ou distance entre deux vecteurs x et y de C^n est :

$$d(x, y) = p(x - y)$$

I.2 - Equivalence entre normes et semi-normes [4]

I.2.1 - Théorème 1

Sur C^n (comme sur tout espace vectoriel de dimension finie) toutes les normes sont équivalentes, c'est-à-dire qu'à tout ensemble de deux normes p et q, on peut faire correspondre deux constantes a et b ($o < a, b < \infty$) telles que :

$$\forall x \in C^n \quad : \quad a\, p(x) \leqslant q(x) \leqslant b\, p(x)$$

I.2.2 - Théorème 2

Si p est une semi-norme sur C^n, q une norme sur C^n, il existe une constante k finie, non nulle telle que :

$$\forall x \in C^n \quad : \quad p(x) \leqslant k\, q(x)$$

Le théorème 1 est classique ; le théorème 2 s'en déduit aisément en considérant :

$$\phi(x) = \text{Max}\ \{p(x),\ q(x)\}$$

l'équivalence de ϕ et de q se traduisant par :

$$\text{Max}\ \{p(x),\ q(x)\} \leqslant K\, q(x) \qquad (k > o)$$

I.2.3 - Théorème 3

Toute semi-norme p (resp. norme) est une application continue, pour toute norme q, de C^n dans R_+.

En effet :

Soit $q(x - y) \leqslant \eta$;

d'après (ii), nous aurons :

$$|p(x) - p(y)| \leqslant p(x-y)$$

et d'après le théorème 2 (resp. Théorème 1) nous aurons :

$$|p(x) - p(y)| \leqslant p(x - y) \leqslant K\, q(x - y)$$

Il suffit de choisir $\eta = \varepsilon / K$ (resp. ε / a)

I.3 - Caractérisation géométrique des normes et semi-normes

I.3.1 - Convexe équilibré

On appellera "Convexe équilibré"[5], tout sous-ensemble fermé $\mathcal{B}$ de C^n, contenant l'origine et ayant les propriétés suivantes :

a) $x, y \in \mathcal{B}$, $0 \leqslant \lambda \leqslant 1 \Longrightarrow \lambda x + (1-\lambda)\, y \in \mathcal{B}$ -Convexité-

b) $x \in \mathcal{B}$, $\omega \in C$: $|\omega| \leqslant 1 \Rightarrow \omega\, x \in \mathcal{B}$ -équilibre-

On désignera par $\alpha\, \mathcal{B}$, l'ensemble

$$\alpha\, \mathcal{B} = \{x \in C^n / \quad x = \alpha\, u \ ; \ u \in \mathcal{B}\}$$

et par $\mathcal{B} + \mathcal{B}'$ l'ensemble

$$\mathcal{B} + \mathcal{B}' = \{x \in C^n / \ x = u + u' \ ; \ u \in \mathcal{B};\ u' \in \mathcal{B}'\}$$

I.3.2 - Théorème 3'

Pour tout convexe équilibré $\mathcal{B}$, l'application $p : C^n \longrightarrow \mathbb{R}_+$

$$p\,(x) = \mathrm{Inf}\ \{\alpha \ / \ \alpha \in \mathbb{R}_+ \ , \quad x \in \alpha\, \mathcal{B}\ \}$$

est une semi-norme sur C^n.

Preuve

Notons que, d'après la définition précédente de la semi-norme p ,nous avons

$$\forall x \in C^n \ : \quad x \in p\,(x)\ \mathcal{B}$$

et que $p\,(x)\ \mathcal{B}$ est le plus petit convexe équilibré contenant x.

1. Soit $\lambda \in C$; posons $\omega = \dfrac{\bar{\lambda}}{|\lambda|}$ $(|\omega| = 1)$

$$y = \lambda\, x \Longrightarrow \omega\, y = |\lambda|\, x$$

et $y \in p\,(y)\mathcal{B} \Longrightarrow |\lambda| x \in p\,(\lambda\, x)\ \mathcal{B}$

d'où $x \in \frac{p(\lambda x)}{|\lambda|} \mathcal{B}$

et ce convexe est le plus petit convexe équilibré contenant x d'où la propriété (i).

2. Soit x et y deux vecteurs de C^n

Pour tout $\nu \in R_+$ tel que $x \in \nu \mathcal{B}$

et pour tout $\mu \in R_+$ tel que $y \in \mu \mathcal{B}$

nous avons

$$x + y \in \nu \mathcal{B} + \mu \mathcal{B} \subset (\nu + \mu)\mathcal{B} \Longrightarrow p(x + y) \leq \nu + \mu$$

En prenant $\nu = p(x)$ et $\mu = p(y)$, nous avons la propriété (ii)

I.3.3- Théorème 4

Lorsque $\mathcal{B}$ est borné, l'application p est une norme

Preuve

Supposons qu'il existe un $x_o \neq 0$ tel que $p(x_o) = o$

alors pour la suite non bornée $x_o, 2x_o, \ldots\ldots 2^n x_o = x_n$

on a d'après i) :

$$p(x_n) = o \qquad \text{d'où} \quad x_n \in \mathcal{B}$$

et cela quelque soit l'indice n; donc $\mathcal{B}$ n'est pas borné, d'où la contradiction.

I.3.4 - Théorème 3'

Pour toute semi-norme p sur C^n, la partie $\mathcal{B}$ de C^n

$$\mathcal{B} = \{x \in C^n \ / \ p(x) \leq 1\}$$

est un convexe équilibré.

Preuve

1) $\mathcal{B}$ est convexe d'après (ii) et il possède la propriété b) d'après i)

2) $\mathcal{B} = \bar{\mathcal{B}}$: image réciproque du segment $[0,1]$ par p continue

3) $\mathcal{B}$ contient l'origine :

Démontrons que $\mathcal{B}$ contient un ouvert contenant l'origine

Pour p donné et pour toute norme q, il existe α tel que

$$\forall x \in C^n \quad : \quad p(x) \leq \alpha \, q(x)$$

ainsi l'ouvert θ de C^n

$$\theta = \{x \in C^n \ / \ q(x) < \frac{1}{\alpha}\}$$

contient l'origine et est contenu dans $\mathcal{B}$.

I.3.5. - Théorème 4'

Si p est une norme sur C^n, la partie $\mathcal{B}$ de C^n

$$\mathcal{B} = \{x \in C^n \ / \ p(x) \leq 1\}$$

est un convexe équilibré borné.

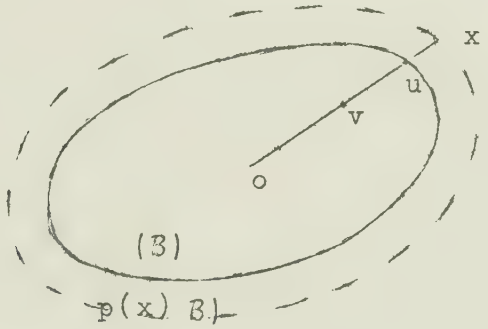
Cette réciproque du théorème 4 est immédiate.

Ainsi nous voyons qu'il revient au même de se donner sur C^n, *une semi-norme p (resp. une norme) ou un domaine convexe équilibré* $\mathcal{B}$ *(resp. borné)*

$\mathcal{B}$ sera appelé boule-unité de la semi-norme p (resp. norme)

Remarque

Dire que $\mathcal{B}$ est borné revient à dire que toute demi-droite issue de l'origine coupe $\mathcal{B}$ en une point u et un seul, à distance finie et que pour ce point $p(u) = 1$.



Nous aurons :

$$p(x) = \frac{\overline{ox}}{\overline{ou}} < \frac{\overline{ox}}{\overline{ov}}$$

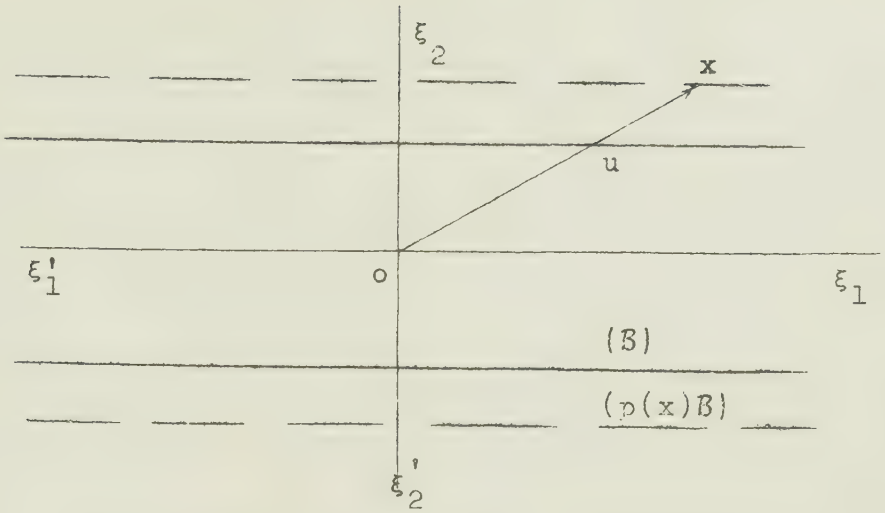
I.3.6 - Exemples sur R^2

Exemple 1 : Considérons l'ensemble $\mathcal{B}$ des points (ξ_1, ξ_2) du plan, compris entre les deux droites D_1 : $\xi_2 = +1$ et D_2 : $\xi_2 = -1$

$$\mathcal{B} = \{x \in R^2 \ / \ |\xi_2| \leq 1\}$$

et nous définirons la semi-norme p par :

$$p(x) = |\xi_2|$$



Remarquons que B n'est pas borné et que tout point x de l'axe ξ_1' ξ_1 est tel que

$$p(x) = |\xi_2| = o$$

Exemple 2

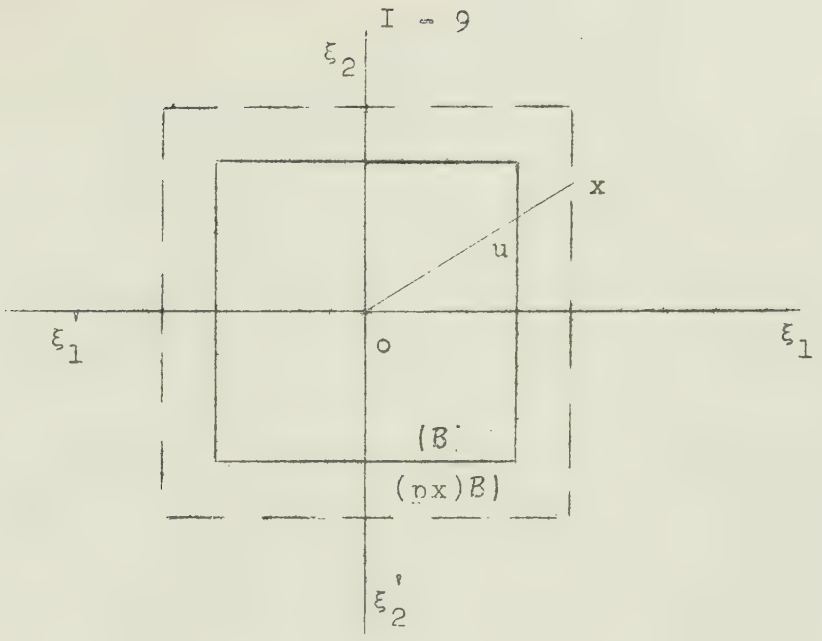
En prenant la partie B du plan compris entre les droites $D_1 : \xi_2 = +1$; $D_2 : \xi_2 = -1$; $D_1' : \xi_1 = +1$ et $D_2' : \xi_1 = -1$

$$B = \{x \in R^2 \ / \ \text{Max} \ (|\xi_1| \ , \ |\xi_2| \ ) \leqslant 1\}$$

Nous définirons la norme p

$$p \ (x) = \text{Max} \ (|\xi_1| \ , \ |\xi_2| \ )$$

B est alors borné et toute droite issue de o coupe B en un point u à distance finie.



I.4. - Norme duale d'une norme p

I.4.1 - Définition

Considérons une norme p sur C^n et $B \subset C^n$ sa boule unité. La partie B' de C^n

$$B' = \{x \;/\; u \in B,\ |u^H x| \leq 1\}$$

est un convexe équilibré borné appelé polaire de B .

Nous pouvons alors caractériser la norme ayant B' pour boule unité ; nous appellerons cette norme duale de p sur C^n et elle sera notée p^*. [5]

La norme p^* sera définie en fonction de p par l'une des relations équivalentes suivantes :

$$p^*(x) = \sup_{p(y)=1} |y^H x| = \sup_{\substack{y \in C^n \\ y \neq o}} \frac{|y^H x|}{p(y)}$$

En effet :

Soit $x \in C^n$; $x \in p^{\#}(x)\ \mathcal{B}' \Longrightarrow \exists y \in \bar{\mathcal{B}}' : x = p^{\#}(x)\ y$

et $\forall u \in \mathcal{B} \quad |x^H u| = p^{\#}(x)\ |y^H u| \Longrightarrow |x^H u| \leq p^{\#}(x)$

d'où

$$p^{\#}(x) \geq \sup_{u \in \mathcal{B}} |x^H u|$$

or $\mathcal{B}$ étant fermé, borné, la fonction continue $|x^H u|$ atteindra ses bornes pour tout point de la frontière de $\mathcal{B}$, c-à-d pour tout $V \in \mathcal{B}$ tel que $p(V) = 1$;

ainsi

$$p^{\#}(x) \geq \sup_{p(V)=1} |x^H V|$$

De plus, l'égalité sera atteinte puisque, pour tout y appartenant à la frontière de $\mathcal{B}'$, il existe un $u_o \in \mathcal{B}$ tel que $|y\ u_o| = 1$.

Notons aussi que tout vecteur $y \in C^n$ pourra s'écrire

$$y = p(y)V \quad \text{avec } p(V) = 1$$

d'où

$$p^{\#}(x) = \sup_{p(V)=1} |x^H V| = \sup_{\substack{y \in C^n \\ y \neq o}} \frac{|x^H y|}{p(y)}$$

et nous en déduisons immédiatement

$$\boxed{\forall x,y \in C^n \ : \ p^*(x)\, p(y) \geq |x^H y|} \quad \text{(inégalité de Holder)}$$

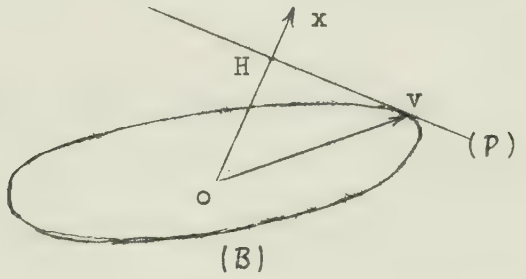
Nous remarquons, de plus, que pour un y donné, nous avons :

$$\forall x \in C^n,\ x \neq o \ : \ p(y) \geq \frac{|x^H y|}{p^*(x)} \Longrightarrow p(y) = \sup_{\substack{x \in C^n \\ x \neq o}} \frac{|x^H y|}{p^*(x)} = \sup_{p(V)=1} |V^H y|$$

et nous en déduisons la propriété réciproque de la dualité

$$(p^*)^*(y) = p(y) \quad \text{et} \quad (\mathcal{B}')' = \mathcal{B}$$

I.4.2 - Interprétation géométrique



Soit x un point quelconque de C^n. Le plan (P) perpendiculaire à ox et s'appuyant en V sur $\mathcal{B}$ coupe ox en H.

On a :

$$p^*(x) = \overline{OH}.\ \overline{OX}$$

Le plan (P) est appelé plan d'appui de $\mathcal{B}$.

I.5 - Principales normes employées en Analyse Numérique

I.5.1 - Normes de Holder

Nous noterons H_p de telles normes que nous définirons sur C^n par :

$$H_p(x) = \left(\sum_i |x_i|^p\right)^{\frac{1}{p}} \quad (1 \leq p \leq \infty)$$

Les normes de Holder les plus courantes seront :

pour $p = \infty$ $\quad H_\infty(x) = \max_i (|x_i|)$

$p = 1$ $\quad H_1(x) = \sum_i |x_i|$

$p = 2$ $\quad H_2(x) = \left(\sum_i |x_i|^2\right)^{\frac{1}{2}} = (x^H x)^{\frac{1}{2}}$ -norme euclidienne-

I.5.2 - Propriété importante de normes de Holder

La duale d'une norme H_p sur C^n est la norme H_q telle que

$$\frac{1}{p} + \frac{1}{q} = 1$$

I.5.3 - Norme ellipsoidale

Soit B une matrice (n, n) définie positive ($B = R^H R$)

l'application $H : C^n \longrightarrow R_+$

$$H(x) = (x^H B x)^{\frac{1}{2}} = H_2(R x)$$

est une norme sur C^n.

La boule unité de cette norme est un ellipsoïde. Les axes de cet ellipsoïde sont les vecteurs propres de B.

La duale de H est caractérisée par :

$$H^{*}(x) = (x^{H} B^{-1} x)^{\frac{1}{2}} = H_2\left[(R^{-1})^{H} x\right]$$

Le polaire de B sera alors un ellipsoïde ayant même axe.

I.5.4 - Décomposition d'une norme [6]

Définition

Une décomposition d'une norme p sur C^n est une application de C^n dans C^n telle que :

$$x \in C^n \quad : \quad p(x) = z^{H}(x).x$$

Exemples:

1) Pour $p = H_1$, on pourra écrire :

$$H_1(x) = z^{H}(x).x$$

avec

$$z(x) = \begin{vmatrix} \frac{|x_1|}{x_1} \\ \vdots \\ \frac{|x_n|}{x_n} \end{vmatrix}$$

Lorsque H_1 est définie sur R^n, z(x) devient

$$z(x) = \begin{vmatrix} \mathrm{Sig}\ (x_1) \\ \vdots \\ \mathrm{Sig}\ (x_n) \end{vmatrix}$$

avec $\mathrm{sig}\ (x_i) = +1 \qquad x_i \geq 0$

$\mathrm{sig}\ (x_i) = -1 \qquad x_i < 0$

2) Pour $p = H_\infty$, nous aurons

$$H_\infty = z^H(x).x$$

avec

$$z(x) = \begin{vmatrix} 0 \\ \vdots \\ \dfrac{|x_i|}{x_i} \\ \vdots \\ 0 \end{vmatrix} \qquad \text{où } |x_i| = \max_k |x_k|$$

3) Pour $p = H_2$, on aura :

$$z(x) = \frac{x}{H_2(x)}$$

et plus généralement pour une norme ellipsoïdale

$$z(x) = \frac{Bx}{(x^H Bx)^{\frac{1}{2}}}$$

4) Pour toute norme de Holder

$$H_p(x) = z^H(x).x$$

avec

$$z(x) = \frac{1}{[H_p(x)]^{p-1}} \left| \begin{array}{c} \frac{|x_1|^p}{x_1} \\ \vdots \\ \frac{|x_n|^p}{x_n} \end{array} \right.$$

Notons qu'ici :

$$H_p^*|z(x)| = H_q|z(x)| = 1 \quad ; \quad H_p(x).H_p^*(z) = z^H x$$

et $z(x)$ est le vecteur culminant fonction de x de la norme duale H_q

$$H_p(x) = \sup_{\substack{z \in C^n \\ z \neq o}} \frac{|x^H z|}{H_q(z)} = |x^H z(x)|$$

I.5.5 - Autres propriétés des normes de Holder

1) Pour toute matrice $A \in M_{(m,n)}$ et pour tout vecteur $x \in C^n$, on a :

$$H_p(Ax) \leq H_p(Ax)\, H_q(x)$$

En effet :

$$H_p(Ax) = \left[\left| \sum_{k=1}^{n} a_{ik}\, x_k \right|^p + \cdots\cdots\cdots + \left| \sum_{k=1}^{n} a_{mk}\, x_k \right|^p \right]^{\frac{1}{p}}$$

$$= \left[|A_{1.}x|^p + \cdots\cdots\cdots + |A_{m.}x|^p \right]^{\frac{1}{p}}$$

où $A_{i.}$ désigne la i^{e} colonne de A $\quad (i = 1 \text{ ----- } m)$.

or sur C^n :

$$|A_{i.}x| \leq H_p(A_{i.}^t)\ H_q(x) \quad (i = 1 \text{ ---- } m)$$

donc

$$H_p(Ax) \leq H_q(x) \left[\sum_{i=1}^{m} H_p(A_{i.}^t) \right]^{\frac{1}{p}}$$

En remarquant que :

$$(1) \qquad H_p(A) = \left[\sum_{i=1}^{m} H_p(A_{i.}^t) \right]^{\frac{1}{p}}$$

nous avons l'inégalité démontrée.

(2) Pour tout $x \in C^n$ et pour tout $y \in C^m$, on a

$$H_p(x\, y^H) = H_p(x).\ H_p(y)$$

Cette égalité est immédiate d'après la relation (1)

TABLEAU RECAPITULATIF sur R^n

	Définitions	Duale	Boule-Unité	Vecteur de décomposition
H_1	$\sum_{i=1}^{n} \lvert x_i \rvert$	H_∞	Parallélotope	$z(x) = \begin{vmatrix} \text{Signe } (x_1) \\ \vdots \\ \text{Signe } (x_n) \end{vmatrix}$
H_2	$(\sum \lvert x_i \rvert^2)^{\frac{1}{2}} = (x^t x)^{\frac{1}{2}}$	H_2	Sphère	$z(x) = \dfrac{x}{\sqrt{x^+ x}}$
H_∞	$\max_i \lvert x_i \rvert$	H_1	Cube	$z(x) = \begin{vmatrix} 0 \\ \vdots \\ \text{sig}(x_i) = \text{sig}(\max_k \lvert x_k \rvert) \\ \vdots \\ 0 \end{vmatrix}$
H_p	$(\sum \lvert x_i \rvert^p)^{\frac{1}{p}}$	$H_q(\frac{1}{p} + \frac{1}{q}) = 1$		$z(x) = \dfrac{1}{[H_p(x)]^{\frac{p}{q}}} \begin{vmatrix} \text{sig}(x_1)\lvert x_1 \rvert^{p-1} \\ \vdots \\ \text{sig}(x_n)\lvert x_n \rvert^{p-1} \end{vmatrix}$
H	$(x^+ Bx)^{\frac{1}{2}}$ $B = R \cdot R^+$	$H^*(x) = (x^t B^{-1} x)^{\frac{1}{2}}$	Ellipsoïde	$z(x) = \dfrac{Bx}{(x^t Bx)^{\frac{1}{2}}}$

BORNES D'EQUIVALENCES

	H_1	H_2	H_∞
H_1	1 ; 1	$\frac{1}{\sqrt{n}}$; 1	$\frac{1}{n}$; 1
H_2	1 ; $\sqrt{n}$	1 ; 1	$\frac{1}{\sqrt{n}}$; 1
H_∞	1 ; n	1 ; $\sqrt{n}$	1 ; 1

II - Norme de matrices

II.1 - Définition

Nous désignerons par $M(m, n)$ l'espace vectoriel à $m \times n$ dimensions des matrices de type (m, n) à coefficient dans C.

Nous pouvons considérer une matrice comme un vecteur de $C^{m \times n}$ et définir une norme de matrice de la même façon que précédemment :

Une application $N : M_{(m,n)} \longrightarrow R^+$ sera une norme de matrice sur $M_{(m,n)}$ si elle possède les propriétés suivantes :

(j) $\lambda \in C$: $N(\lambda A) = |\lambda| N(A)$ [en particulier $N(o) = o$]

(jj) $N(A + B) \leq N(A) + N(B)$

(jjj) $N(A) = o \quad A = o$ [ou avec (j) $N(A) = o \Leftrightarrow A = o$]

Mais une matrice représente aussi une application linéaire et il peut être intéressant de ne pas se contenter de ce premier aspect et de faire intervenir la propriété supplémentaire relative au produit de deux opérateurs linéaires.

C'est la raison pour laquelle nous introduirons la notion de norme multiplicative.[7]

Une norme M sera dite multiplicative, si en plus des trois propriétés précédentes, elle satisfait à la propriété :

(jjjj) $M(A.B) \leq M(A).\ M(B)$

II.2 - Normes de matrices associées à des normes de vecteurs

II.2.1 - Définition

Considérons une norme p définie sur C^n et une norme q définie sur C^m

Pour $A \in M_{(m,n)}$, posons :

$$\boxed{Spq(A) = \sup_{\substack{x \in C^m \\ x \neq o}} \frac{p(Ax)}{q(x)}}$$

alors Spq est une norme sur $M_{(m,n)}$

Les 3 axiomes de normes se démontrent aisément [2]

II.2.2 - Caractérisation géométrique et propriétés des normes "Spp"

Soit B un convexe équilibré borné ($B \subset C^n$). Nous désignerons par $A\,B$, l'ensemble

$$A\,B = \{y \in C^n / y = Ax \ ; \ x \in B\}$$

où A est une matrice carrée (n,n)

Théorème 5

L'application $S : M_{(n,n)} \longrightarrow R_+$ telle que

$$S(A) = \mathrm{Inf}\left[\nu \ ; \ \nu > o \ , \ A\,B \in \nu\,B\right]$$

est une norme multiplicative sur $M_{(n,n)}$

Preuve :

1) Notons d'abord qu'ici aussi, nous aurons

$$A \in M_{(n,n)} \quad ; \quad A\,B \subset \gamma(A)\,B$$

Soit $\lambda \in \mathbb{C}$.

Posons $B = \lambda A$ et $\omega = \bar{\lambda} / |\lambda|$; $|\omega| = 1$

nous aurons

$$\omega B = |\lambda| A$$

et puisque B est un convexe équilibré

$$B B \subset \gamma(B) B \Longrightarrow \lambda A B \subset \gamma(B) B$$

d'où

$$A \in \frac{\gamma(B)}{|\lambda|} B$$ et ce convexe est le plus petit convexe équilibré contenant A ; d'où

$$|\lambda|\, \gamma(A) = \gamma(\lambda A)$$

2) Soient A et B deux matrices de $M_{(n,n)}$

Pour tout $\alpha \in \mathbb{R}_+$: $B B \subset \alpha B$ et pour tout $\gamma \in \mathbb{R}_+$: $BB \subset \gamma B$

nous avons

$$(A + B) B \subset A B + BB = \alpha B + \gamma B = \{x \in \mathbb{C}^n / x = \gamma u + \gamma v ; u, v \in B\}$$

et en considérant la norme p associée à B

$$\forall x \in \alpha B + \gamma B \quad : \quad p(x) = p(\alpha u + \gamma V) \leq \alpha + \gamma$$

Ainsi

$$x \in p(x)\ \mathcal{B} \subset (\alpha + \gamma)\ \mathcal{B} \Longrightarrow \alpha\ \mathcal{B} + \gamma\ \mathcal{B} \subset (\alpha + \gamma)\ \mathcal{B}$$

et comme il est évident que :

$$(\alpha + \gamma)\ \mathcal{B} \subset \alpha\ \mathcal{B} + \gamma\ \mathcal{B}$$

nous aurons

$$(A + B)\mathcal{B} \subset \alpha\ \mathcal{B} + \gamma\ \mathcal{B} = (\alpha + \gamma)\mathcal{B}$$

En choisissant

$\alpha = S(A)$; $\gamma = S(B)$, nous aurons

$$(A + B)\ \mathcal{B} \subset \left[S(A) + S(B)\right]\mathcal{B} \Longrightarrow S(A+B) \leq S(A) + S(B)$$

d'où la propriété jj)

3) La propriété (jjj) est évidente.

4) Démontrons que la norme S est multiplicative

Pour tout $\alpha \in R^+$: $A\mathcal{B} \subset \alpha\ \mathcal{B}$

et pour tout $\gamma \in R^+$: $B\ \mathcal{B} \subset \gamma\ \mathcal{B}$

Nous avons en posant $\alpha = S(A)$; $\beta = S(B)$

$$(AB)\ \mathcal{B} = A\ (B\ \mathcal{B}) \subset A(\gamma\ \mathcal{B}) = \gamma(A\ \mathcal{B}) \subset \gamma\alpha\ \mathcal{B} \Longrightarrow S(AB) \leq S(A)\ S(B)$$

Nous voyons ici aussi, que tout convexe équilibré, borné, définit sur $M_{(m,n)}$ *une norme de matrice multiplicative.*

II.2.3 - Lemme 1

La norme de vecteur p et la norme S de matrice, définies par le même convexe équilibré $\mathcal{B}$ sont telles

$$\forall A \in M_{(m,n)} : \quad S(A) = pp(A)$$

En effet

1. $\forall x \in \mathbb{C}^n : x \in p(x)\mathcal{B} \Longrightarrow Ax \in p(x)\, A\mathcal{B} \subset p(x)\, S(A)\mathcal{B} \Rightarrow p(Ax) \leq p(x) S(A)$

d'où

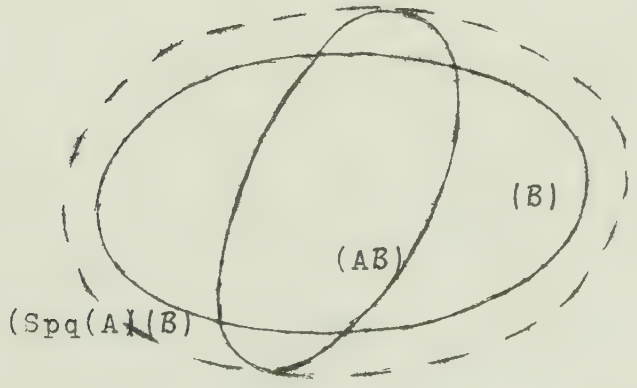
$$S(A) \geq \sup_{\substack{x \in \mathbb{C}^n \\ x \neq o}} \frac{p(Ax)}{p(x)} = Spp(A)$$

2. Soit $x \in A\mathcal{B}$: $x = Au$; $u \in \mathcal{B}$ $(p(u) \leq 1)$

$p(Au) \leq Spp(A)\, p(u) \leq Spp(A) \Longrightarrow x \in p(x)\mathcal{B} = p(Au)\mathcal{B} \subset Spp(A)\, \mathcal{B}$

d'où

$$A\mathcal{B} \subset Spp(A)\, \mathcal{B} \Longrightarrow S(A) \leq Spp(A)$$



II.2.4 - Une propriété des normes spp

Etant donné $A \in M_{(m,n)}$, pour toute norme p de vecteur sur C^n, nous avons

$$Spp\ (A) = Sp^{*}p^{*}\ (A^{H})$$

En effet

$$\forall\ x,y \in C^n \quad : |x^H A\ y| \leq p^{*}(x)\ p(Ay) \leq p^{*}(x)\ Spp(A)\ p(y)$$

Pour une matrice A fixée, il existe un y_1 tel que

$$p(Ay_1) = Spp\ (A)\ p\ (y_1)$$

et pour cet y_1, il existe un x_1 tel que

$$|x_1^H\ Ay_1| = p^{*}(x_1)\ p\ (Ay_1)$$

d'où

$$|x_1^H A\ y_1| = p^{*}\ (x_1)\ \ p(y_1)\ Spp(A)$$

de plus

$$|x_1^H A\ y_1| = |y_1^H A^H\ x_1| \leq p(y_1)\ p^{*}(A^H x_1) \leq p^{*}(x_1)p(y_1)Sp^{*}p^{*}(A^H)$$

$$\text{donc} \qquad Spp(A) \leq S_{p^{*}\ p^{*}}(A^H)$$

De même, il existe x_2 et y_2 tels que :

$$|x_2^H A\ y_2| = p^{*}(x_2)\ p\ (y_2)\ Sp^{*}p^{*}\ (A^H) \Longrightarrow S_{p^{*}\ p^{*}}\ (A^H) \leq Spp(A)$$

donc

$$Spp\ (A) = S_{p^{*}\ p^{*}}\ (A^H)$$

II.2.5 - Lemme 2

Pour toute norme multiplicative M, il existe des normes de vecteurs p_M telles que

$$\forall\, A \in M_{(m,n)} \quad : \quad \rho(A) \leq S_{p_M\, p_M}(A) \leq M(A)$$

où $\rho(A)$ = module de la plus grande valeur propre de A.

1. Soit b un vecteur non nul de C^n

La fonction $P_M(x) = M(x\, b^H)$ est une norme sur C^n (évident)

et $\forall x : P_M(Ax) = M(A\, x\, b^H) \leq M(A)\, M(x\, b^H) = M(A)\, P_M(x)$

d'où $M(A) \leq S_{p_M p_M}(A)$

autrement dit la norme $S_{p_M\, p_M}$ est plus "serrée" que M.

Nous voyons ainsi qu'il existe pour chaque vecteur $b \neq o$ de C^n une norme correspondante p_M ; dans la pratique, on prendra pour b, le premier vecteur de la base fondamentale de C^n soit e_1.

$$x\, b^H = x\, e_1^H = X = \begin{vmatrix} x_1 & o & \cdots & o \\ \vdots & & & \vdots \\ x_n & o & \cdots & o \end{vmatrix}$$

2. Soit $A\, x = \lambda\, x \quad (x \neq o)$

$$p_M(A\, x) = |\lambda| p(x) \Longrightarrow |\lambda| \leq S_{p_M p_M}(A)$$

II.3 - Normes de matrices employées en Analyse Numérique

II.3.1 - Normes de Holder

L'ensemble $M_{(m,n)}$ étant isomorphe (algébriquement) à $C^{m\times n}$ on pourra définir une norme sur $M_{(m,n)}$ comme nous l'avons fait sur $C^{m\times n}$, par exemple nous pouvons définir ici aussi une norme de matrice de Holder d'ordre p notée H_p.

$$H_p(A) = \left(\sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}|^p \right)^{\frac{1}{p}}$$

$$(1 \leqslant p \leqslant \infty)$$

Une telle norme sera appelée norme (de matrice) de Holder d'ordre p.

Nous aurons ici aussi :

Pour $p = \infty$ $\quad H_\infty(A) = \underset{i,j}{\text{Max}} |a_{ij}|$

$p = 1$ $\quad H_1(A) = \sum_i \sum_j |a_{ij}|$

$p = 2$ $\quad H_2(A) = \left(\sum_i \sum_j |a_{ij}|^2 \right)^{\frac{1}{2}} = \left[\text{trace } (A^H A) \right]^{\frac{1}{2}}$

II.3.2. - Normes Spp

Les normes Spp les plus employées sont :

$$S_{H_1 H_1}(A) = \underset{\substack{x \in C^m \\ x \neq o}}{\text{Sup}} \frac{H_1(Ax)}{H_1(x)} = \underset{j}{\text{Sup}} \sum_{i=1}^{m} |a_{ij}| \quad \text{(notée } S_{11}(A)\text{)}$$

$$S_{H_\infty H_\infty}(A) = \sup_{\substack{x \in C^m \\ x \neq o}} \frac{H_\infty(Ax)}{H_\infty(x)} = \sup_i \sum_j |a_{ij}| \qquad \text{(notée } S_{\infty\infty}(A)\text{)}$$

$$S_{H_2 H_2}(A) = \sup_{\substack{x \in C^m \\ x \neq o}} \frac{H_2(Ax)}{H_2(x)} = \lambda_M \qquad \text{(notée } S_{22}(A)\text{)}$$

λ_M plus grande valeur propre de $A^H A$.

La 3ème relation se démontre facilement en décomposant chaque x de C^n suivant les n vecteurs propres de $A^H A$.

Pour démontrer les deux premières relations, nous pouvons démontrer d'abord que :

Pour toute norme p définie sur C^n

$$S_{H_\infty p}(A) = \sup_i \; p^*(A^H_{.i})$$

où $A_{.i}$: i ème ligne de A.

Preuve

Nous avons $$S_{H_\infty p}(A) = \sup_{\substack{x \in C^m \\ x \neq o}} \frac{\sup_i |A_{.i} x|}{p(x)}$$

$$\forall\, x \in C^n : |A_{.i}\, x| \leq p^*(A^H_{.i})\; p(x) \Longrightarrow \sup_i |A^H_i\, x| \leq \sup_i \; p^*(A^H_i)\; p(x)$$

d'où

$$\operatorname*{Sup}_{\substack{x \in C^m \\ x \neq o}} \frac{\operatorname*{Sup}_i |A_i^H x|}{p(x)} \leq \operatorname*{Sup}_i p^*(A_i^H)$$

L'égalité sera atteinte ; en effet :

soit k tel que $p^*(A_{.k}^H) = \operatorname*{Sup}_i p^*(A_i^H)$

et il existe x_o tel que

$$|A_{.k}^H x_o| = p^*(A_k^H)\, p(x_o)$$

Remarquons de plus que nous avons

$$S_{\infty 1}(A) = \operatorname*{Max}_i (\operatorname*{Max}_j |a_{ij}|) = \operatorname*{Max}_{ij} |a_{ij}| = H_\infty(A)$$

III - Normes Vectorielles de Vecteurs [9]

III.1 - Définition

Une application $p : C^n \longrightarrow R_+^k$ est une norme vectorielle d'ordre k si elle vérifie les axiomes suivants :

(L) $\forall \lambda \in C : p(\lambda x) = |\lambda| p(x)$

(LL) $\forall x,y \in C^n : p(x+y) \leq p(x+y)$ (composante à composante)

(LLL) $p(x) = o \Longrightarrow x = o$ (ou avec (L) $p(x) = o \Longleftrightarrow x=o$)

Nous noterons p_i ($i = 1.....k$), les k composantes de p ; Les applications $p_i : C^n \longrightarrow R_+$ sont des semi-normes sur C^n ; en effet $p_i(x) = o$ n'entraine pas nécessairement $x = o$.

Exemple :

Norme vectorielle d'ordre n définie par :

$p_i(x) = |x_i|$ (module de la ie composante de x)

III.2 - Caractérisation géométrique

III.2.1 - Théorème 6

Soient k semi-normes $p_1 \ldots\ldots p_k$ sur C^n, et $\mathcal{B}_1 \ldots\ldots \mathcal{B}_k$ leur boule unité telles que :

$$\bigcap_{i=1}^{k} \mathcal{B}_i \quad \text{borné}$$

alors l'application $p : C^n \longrightarrow R_+^k$

$$x \in C^n \longrightarrow p(x) = \begin{vmatrix} p_1(x) \\ \vdots \\ p_k(x) \end{vmatrix}$$

est une norme vectorielle.

Preuve

1. Les axiomes (L) et (LL) se satisfont aisément.

2. Si $p(x) = o \not\Longrightarrow x = o$, il existerait $x_o \neq o$ tel que $p(x_o) = o$; alors la suite non bornée $x_o, 2x_o, \text{-------}, 2^n x_o = x_n$ donnerait :

$$\forall_n \quad : \quad p(x_n) = o \implies x_n \in \mathcal{B}_i$$

et ainsi l'intersection des $\mathcal{B}_i$ ne serait pas bornée d'où contradiction, d'où (LLL) démontré.

Intersement

III.2.2 - Théorème 6'

Les composantes p_i de toute norme vectorielle p d'ordre k sur C^n, sont des semi-normes et le convexe $\bigcap_{i=1}^{k} \mathcal{B}_i$ est borné.

Preuve

Les axiomes (L) et (LL) entrainent les axiomes (i) et (ii) pour les p_i et de plus $p_i(x) = o \not\Longleftrightarrow x = o$

Supposons alors que $\bigcap_{i=1}^{k} \mathcal{B}_i$ ne soit pas borné. Il existerait alors une suite $x_1, \text{------} x_n$ de C^n, appartenant à $\bigcap_{i=1}^{k} \mathcal{B}_i$ tel que pour toute norme q de C^n, on ait

$$\lim_{n \to \infty} q(x_n) = \infty$$

or $\operatorname{Max}_i \{p_i(x)\}$ est une norme sur C^n

d'où

$$\lim_{x_n \longrightarrow \infty} \operatorname{Max}_i \{p_i(x_n)\} = \infty \Longrightarrow \exists x_j : p_i(x_j) > 1 \quad (i=1 \text{---} k)$$

or $x_j \in \bigcap_{i=1}^{k} B_i \Longrightarrow p_i(x_j) \leq 1$ d'où contradiction.

Ainsi :

La donnée de k convexes équilibrés à intersection bornée permet de définir une norme vectorielle d'ordre k sur C^n.

III.3 - Normes vectorielles sur une somme de sous-espaces

II.3.1 - Soit C^n décomposé en une somme de r sous-espaces $C^{\alpha i}$

$$C^n = C^{\alpha 1 + \alpha 2 \text{----------} + \alpha r}$$

et $q_1, \text{-------} q_r$ des normes sur ces sous-espaces.

$X \in C^n$ s'écrit $X = \begin{vmatrix} x_1 \\ \vdots \\ x_r \end{vmatrix}$ où $x_j \in C^{\alpha j}$

Alors l'application $q : C^n \longrightarrow R_+^r$

$$x \in C^n \longrightarrow q(x) = \begin{vmatrix} q_1(x_1) \\ \vdots \\ q_r(x_r) \end{vmatrix}$$

est une norme vectorielle d'ordre r sur C^n.

Soient de même

$$C^m = C^{\beta 1} + \text{-------} + \beta_s$$

et p_1 ------- p_s des normes sur les sous-espaces $C^{\beta 1}$ ----- $C^{\beta s}$

$$Y \in C^m \text{ s'écrit } Y = \begin{vmatrix} y_1 \\ \vdots \\ y_s \end{vmatrix} \text{ où } y_i \in C^{\beta j}$$

et l'application $p : C^m \longrightarrow R_+^s$

$$Y \in C^m \longrightarrow p(Y) = \begin{vmatrix} p_1 (y_1) \\ \vdots \\ p_r (y_s) \end{vmatrix}$$

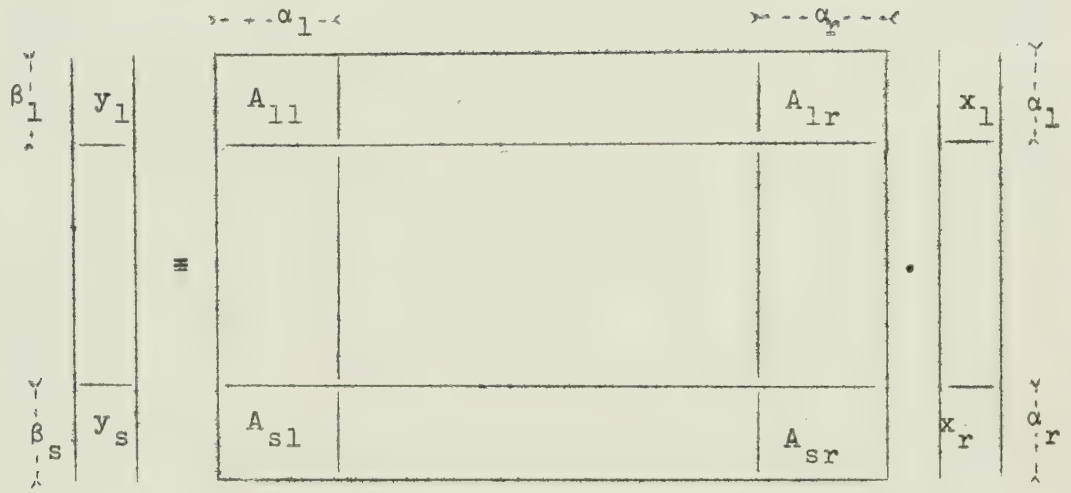
est une norme vectorielle d'ordre s sur C^m.

III.3.2 - <u>Application importante</u>

Soit une matrice $A : C^m \longrightarrow C^n$ décomposée en matrice-bloc A_{ij} de type (β_i , α_j) :

$$Y = A X \longrightarrow (2) \quad y_i = \sum_{j=1}^{r} A_{ij} x_j \quad (i=1\text{-----} s)$$

Schématiquement, on a :



On démontre qu'il existe des matrices M de type (s p) à éléments non négatifs telles que :

$$(3) \qquad \forall x \in C^n \quad : \quad p(AX) \leq M.\, q(x)$$

Théorème 7

$$\text{La matrice } M(A) = \left(S_{p_i q_j}(A_{ij}) \right) = \sup_{\substack{x_j \in C^{\alpha_j} \\ x_j \neq 0}} \frac{p_i(A_{ij}\, x_j)}{q_j(x_j)}$$

est la matrice minimale unique vérifiant (3).

En effet :

1. (2) donne

$$p_i\,(y_i) \leq \sum_{j=1}^{s} p_i\,(A_{ij}\,x_j) < \sum_{j=1}^{s} S_{p_i q_j}\,(A_{ij}).q_j\,(x_j) \quad (i=1 \text{---} r)$$

d'où M(A) vérifie (3)

2. Soit

$N = (n_{ij})$ une matrice vérifiant (3) ;

nous aurons

$$p_i\,(y_i) \leq \sum_{j=1}^{s} n_{ij}\,q_j\,(x_j) \qquad (i = 1 \text{-----} r)$$

$$p_i\,\Big(\sum_{j=1}^{s} A_{ij}\,x_j\Big) \leq \sum_{j=1}^{s} n_{ij}\,q_j\,(x_j) \qquad (i=1 \text{----} s)$$

et cela quelque soit x_1 ------ x_s ;

En particulier pour

$$x_k = o \quad (k = 1 \text{----} j-1,\ j+1 \text{----} s)$$

et $x_j \neq o$

d'où

$$p_i\,(A_{ij}\,x_j) \leq n_{ij}\,q_j\,(x_j) \Longrightarrow n_{ij} \geq \frac{p_i\,(A_{ij}\,x_j)}{q_j\,(x_j)} \quad (\forall x_j \in C^{\alpha_j})$$

et

$$n_{ij} \geq \operatorname*{Sup}_{\substack{x_j \in C^{\alpha_j} \\ x_j \neq o}} \frac{p_i(A_{ij}\,x_j)}{q_j\,(x_j)} = S_{p_i q_j}\,(A_{ij})$$

donc M(A) est bien unique.

Chapitre II

Quelques conditions de compatibilité d'un résultat avec les données d'un système linéaire à coefficients incertains.

I - QUELQUES LEMMES

I.1 - Lemme 1 (théorème II.1)

Le lieu des vecteurs $A\tilde{x}$ [$\tilde{x}$ vecteur donné de C^m ; $A \in \mathcal{A}$ - $\mathcal{A}$ convexe équilibré de C^{mxn}] est un convexe équilibré de C^m

Ce résultat, simple et classique nous permet d'écrire

$$N_1(A) \leq \alpha \Rightarrow N_{2(1)}(A\tilde{x}) \leq \alpha.k$$

Les fonctions N_2 et k dépendant de la nature de la norme N_1 correspondent à $\mathcal{A}$

I.2 - Lemme 2 (théorème II.2)

Pour tout vecteur donné $\tilde{x}$ de C^n, le lieu des vecteurs y de C^m

$$y = A\tilde{x},\ S_{pq}(A) \leq \alpha$$

est la boule $\mathcal{B}$, correspondant à la norme p , de rayon $\alpha_q(\tilde{x})$

$$\mathcal{B} = \{y \in C^m / p(y) \leq \alpha_q(\tilde{x})\}$$

Preuve :

1. $y = A\tilde{x}$; $S_{pq}(A) \leq \alpha \Longrightarrow p(y) = p(A\tilde{x}) \leq S_{pq}(A)\, q(\tilde{x}) \leq \alpha\, q(\tilde{x}) \Longrightarrow y \in B$

2. Montrons réciproquement que nous pouvons faire correspondre à tout vecteur $y \in B$, une matrice $\tilde{A}$ (antiscalaire) telle que

$$y = \tilde{A}\,\tilde{x} \quad \text{et} \quad S_{pq}(\tilde{A}) \leq \alpha,$$

Posons :

$$\tilde{A}(y) = \frac{y.\tilde{z}^H(\tilde{x})}{\tilde{z}^H(\tilde{x})\tilde{x}} \qquad \left[x \neq o \Rightarrow q(\tilde{x}) \neq o \Rightarrow \tilde{z}^H\tilde{x} \neq o \Longrightarrow \text{existence de } \tilde{A}\right]$$

$\tilde{z}(\tilde{x})$ étant un vecteur maximisant de $\tilde{x}$ relativement à la norme $q(\tilde{x})$ c'est-à-dire tel que :

$$q(\tilde{x}) = \sup_{\substack{z \in C^n \\ z \neq o}} \frac{|x^H z|}{q^*(z)} = \frac{|x^H \tilde{z}|}{q^*(\tilde{z})}$$

Il est, en effet, clair que $y = \tilde{A}\tilde{x}$, et que

$$\sup_{\substack{x \in C^n \\ x \neq o}} \frac{p(\tilde{A}x)}{q(x)} = \sup_{\substack{x \in C^n \\ x \neq o}} \frac{p(y\tilde{z}^H x)}{|\tilde{z}^H\tilde{x}|\, q(x)} = \frac{p(y)}{|\tilde{z}^H\tilde{x}|} \sup_{\substack{x \in C^n \\ x \neq o}} \frac{|\tilde{z}^H x|}{q(x)}$$

$$= \frac{p(y)}{|\tilde{z}^H\,\tilde{x}|}\, q^*(\tilde{z}) = \frac{p(y)}{q(\tilde{x})} \leq \alpha \Longrightarrow S_{pq}(\tilde{A}) \leq \alpha$$

I.3 - Lemme 2' (Théorême II.3)

Pour un vecteur donné $\tilde{x}$ de C^n, le lieu des vecteurs y de C^m

$$y = A\tilde{x} \; ; \qquad H_p(A) \leq \alpha$$

est la boule $\mathcal{B}$ correspondant à la norme de Holder H_p définie sur C^m, de rayon $\alpha H_q(\tilde{x})$

$$\mathcal{B} = \{y \in C^m \; / \; H_p(y) \leq \alpha \; H_q(\tilde{x})\}$$

La démonstration est analogue à la précédente ; en effet :

1. Soit $y = A\tilde{x} \; ; \qquad H_p(A) \leq \alpha,$

D'après la propriété de H_p du chapitre 1(§ I.5.5), nous avons :

$$H_p(y) = H_p(A\tilde{x}) \leq H_p(A) \; H_q(\tilde{x}) \leq \alpha \; H_q(\tilde{x}) \Longrightarrow y \in \mathcal{B}$$

2. Soit $y \in \mathcal{B} \; ; \qquad H_p(y) \leq \alpha \; H_q(\tilde{x})$

On prendra ici la même matrice $\tilde{A}$

$$\tilde{A} = \frac{y \; \tilde{z}^H}{\tilde{z}^H \; \tilde{x}}$$

avec $\tilde{z}$ tel que : $H_q(\tilde{x}) = \underset{\substack{z \in C^n \\ z \neq o}}{\text{Sup}} \; \frac{|\tilde{x}^H \; z|}{H_p(z)} = \frac{|\tilde{x}^H \; \tilde{z}|}{H_p(\tilde{z})}$

On a d'après § I.5.5. chapitre I

$$H_p(\tilde{A}) = \frac{H_p(y\ z^H)}{|z^H\ x|} = \frac{H_p(y) \cdot H_p(z)}{|z^H\ x|}$$

d'où

$$H_p(\tilde{A}) = \frac{H_p(y)}{H_q(x)} \leq \alpha$$

I.4 - Remarque

Lorsque $x \in C^n$ et $y \in C^m$ sont deux vecteurs donnés, la matrice $\tilde{A}$ antiscalaire

$$\tilde{A} = \frac{y \cdot z^H}{z^H \cdot x}$$

est une matrice minimale pour la norme S_{pq} parmi toutes les matrices A telle que $y = A\,x$

$$S_{pq}(\tilde{A}) = \underset{[A : y = Ax]}{\text{Inf}} \{S_{pq}(A)\} = \frac{p(y)}{q(x)}$$

En effet

Pour A telle que : $y = Ax$, nous avons :

$$\frac{p(\tilde{y})}{q(\tilde{x})} = \frac{p(A\tilde{x})}{q(\tilde{x})} \leq S_{pq}(A) \Longrightarrow \frac{p(\tilde{y})}{q(\tilde{x})} \leq \underset{[A:\tilde{y}=A\tilde{x}]}{\text{Inf}} \{S_{pq}(A)\}$$

et l'égalité est atteinte pour la matrice $\tilde{A}$ puisque d'après ce qui précède

$$Spq(\tilde{A}) = \frac{p(\tilde{y})}{q(\tilde{x})}$$

La même remarque pourra s'appliquer aux normes de Holder et nous aurons

$$H_p(\tilde{A}) = \underset{[A:\tilde{y}=A\tilde{x}]}{\text{Inf}} \{H_p(A)\} = \frac{H_p(\tilde{y})}{H_q(\tilde{x})}$$

I.5 - Lemme 3 (Théorème II.3)

Soient :

p une norme vectorielle d'ordre s sur C^m (de composante P_i)

q une norme vectorielle d'ordre r sur C^n (de composante q_j)

$$M(A) = (m_{ij}) = \underset{\substack{x \in C^{\alpha_j} \\ x_j \neq o}}{\text{Sup}} \frac{P_i(A_{ij}x_j)}{q_j(x_j)} \qquad \begin{matrix} i = 1, \ldots, s \\ j = 1, \ldots, r \end{matrix}$$

$S = (s_{ij} \geq o)$ matrice donnée

alors

Pour un vecteur $\tilde{X}$ donné de C^n, l'ensemble des vecteurs Y de C^m tels que

$$Y = A\tilde{X}; \quad M(A) \leq S \quad \text{(coef. à coef.)}$$

est caractérisé par

$$p(Y) \leq S.q(\tilde{X}) \quad \text{(inégalité vectorielle)}$$

Preuve :

1. Soit $Y \in C^m$ tel que $Y = A\tilde{X}$, $M(A) \leq S$

Nous savons d'après le théorème I.4 que la matrice M(A) est la matrice minimale vérifiant :

$$\forall X \in C^n \; : \; p(AX) \leq M(A) \cdot q(X)$$

d'où

$$P(Y) = p(A\tilde{X}) \leq M.q(\tilde{X}) \Longrightarrow P(Y) \leq S\, q(\tilde{X})$$

2. Nous décomposons les espaces C^m et C^n comme nous l'avons fait au chapitre I § III

$$C^n = C^{\alpha_1+\cdots+\alpha_r} \; ; \; C^m = C^{\beta_1+\cdots+\beta_s}$$

Posons :

$$\tilde{z} \in C^n \; ; \; \tilde{z} = \begin{vmatrix} \tilde{z}_1 \\ \vdots \\ \tilde{z}_r \end{vmatrix}$$

avec $\tilde{z}_j \in C^{\alpha j}$ tel que

$$q_j(\tilde{x}_j) = \sup_{\substack{z_j \in C^{\alpha j} \\ z_j \neq o}} \frac{\left|\tilde{x}_j^H \; z_j\right|}{q_j^*(z_j)} = \frac{\left|\tilde{x}_j^H \cdot \tilde{z}_j\right|}{q_j^*(\tilde{z}_j)}$$

$$\tilde{Z}_j \in C^n \quad ; \quad \tilde{Z}_j = \begin{vmatrix} s_{j1}\, q_1(\tilde{x}_1)\, \dfrac{\tilde{z}_1}{\left|\tilde{z}_1^H \; \tilde{x}_1\right|} = \tilde{Z}_{j1} \in C^{\alpha 1} \\ \vdots \\ s_{jr}\, q_r(\tilde{x}_r)\, \dfrac{\tilde{z}_r}{\left|\tilde{z}_r^H \; \tilde{x}_r\right|} = \tilde{Z}_{jr} \in C^{\alpha r} \end{vmatrix}$$

Notons que

$$(1) \quad \left|\tilde{Z}_j^H \; \tilde{X}\right| = \sum_{k=1}^{r} s_{jk}\, q_k(\tilde{x}_k)$$

et

$$(2) \quad q_k^*(\tilde{Z}_{jk}) = s_{jk} \qquad \begin{vmatrix} k = 1 \text{-----} r \\ j = 1 \text{-----} s \end{vmatrix}$$

Soit alors

$$A = \begin{array}{|c|c|c|} \hline A_{11} & & A_{1r} \\ \hline & & \\ & & \\ \hline A_{s1} & & A_{sr} \\ \hline \end{array} \quad \begin{matrix} \alpha_1 & & \alpha_r \\ & & \\ \beta_1 & & \\ & & \\ \beta_s & & \end{matrix}$$

avec $$A_{ij} = \frac{y_i \, \tilde{Z}_{ij}^H}{\tilde{Z}_i^H \, \tilde{X}}$$

Nous avons

$$\sum_{k=1}^{r} A_{ik} \tilde{x}_k = \sum_{k=1}^{r} \frac{y_i \left| \tilde{Z}_{ik}^H \tilde{x}_k \right|}{\left| \tilde{Z}_i^H \tilde{X} \right|} = y_i \frac{\sum_{k=1}^{r} \left| \tilde{Z}_{ik}^H \tilde{x}_k \right|}{\left| \tilde{Z}_i^H \tilde{X} \right|} = y_i \Longrightarrow Y = A \tilde{X}$$

De plus

$$(3) \qquad \sup_{\substack{x_j \in C^{\alpha_j} \\ x_j \neq o}} \frac{p_i (A_{ij} x_j)}{q_j (x_j)} = \frac{p_i (y_i)}{\left| \tilde{Z}_i^H \tilde{X} \right|} \sup_{\substack{x_j \in C^{\alpha_j} \\ x_j \neq o}} \frac{\left| \tilde{Z}_{ij}^H x_j \right|}{q_j (x_j)} = \frac{p_i (y_i)}{\left| \tilde{Z}_i^H \tilde{X} \right|} q_j^{\#} (\tilde{Z}_{ij})$$

or

$$p(Y) \leq S.\ q(\tilde{X}) \Longrightarrow p_i(y_i) \leq \sum_{k=1}^{r} s_{ik}\, q_k(\tilde{x}_k)$$

$$i = 1 \text{------} s$$

et d'après (1)

$$p_i(y_i) \leq |\tilde{Z}_i^H \tilde{X}|$$

nous donne alors

$$\sup_{\substack{x_j \in C^{\alpha j} \\ x_j \neq o}} \frac{p_i(A_{ij}\, x_j)}{q_j(x_j)} \leq q_j^{*}(\tilde{Z}_{ij}) = s_{ij}$$

et cela quels que soient les indices i et j

donc

$$\forall Y : p(Y) \leq S\ q(\tilde{X}) \Longrightarrow \exists\, A : Y = A\tilde{X};\quad M(A) \leq S$$

II - APPLICATIONS A L'EXISTENCE D'UNE MATRICE SINGULIERE INDISCERNABLE DE A_o ET A LA DISTANCE D'UNE MATRICE AU LIEU DES MATRICES SINGULIERES.

Il est bon de déterminer si le lieu des matrices (carrées) indiscernables d'une matrice donnée A_o contient une matrice singulière.

Les lemmes précédents vont nous permettre de donner des conditions nécessaires et suffisantes d'existence d'une telle matrice.

II.1 Théorème II.4

Pour que le lieu Λ_{A_o} des matrices carrées $\tilde{A}$, indiscernables de A_o -au sens d'une norme S_{pq}, $S_{pq}(\tilde{A} - A_o) \leq \alpha$ - contienne une matrice singulière, il faut et il suffit que :

$$S_{qp}(\Lambda_o^{-1}) \geq \alpha^{-1}$$

Preuve :

D'après le théorème II.2, pour tout vecteur donné x de $C^{\overline{n}}$, l'ensemble des vecteurs y de C^n tels que :

$$y = \tilde{A}x \; ; \qquad S_{pq}(\tilde{A} - A_o) \leq \alpha$$

est la boule $\mathcal{B}_x$ de centre $A_o x$ et de rayon $\alpha q(x)$.

Or, pour qu'il existe une matrice singulière appartenant au lieu des matrices indiscernables de A_o il faut et il suffit qu'il existe une boule B_x contenant le vecteur nul ($\tilde{A}$ singulière $\Longleftrightarrow \exists x \neq o : \tilde{A}\, x = o$) donc il faut et il suffit qu'il existe $x \neq o$ tel que

$$p\,(A_o\, x) \leq \alpha\, q\,(x)$$

et

$$\operatorname*{Inf}_{\substack{x \in C^n \\ x \neq o}} \frac{p\,(A_o\, x)}{q\,(x)} = \frac{1}{\operatorname*{Sup}_{\substack{x \in C^n \\ x \neq o}} \dfrac{q\,(x)}{p\,(A_o x)}} = \frac{1}{\operatorname*{Sup}_{\substack{y \in C^n \\ y \neq o}} \dfrac{q\,(A_o^{-1}\, y)}{p\,(y)}}$$

$$= \frac{1}{S_{qp}\,(A_o^{-1})} \leq \alpha \Longleftrightarrow \exists\, \tilde{A} \text{ singulière} \in A_{A_o}$$

Ainsi la boule de centre A_o et de rayon $\left[S_{qp}\,(A_o^{-1})\right]^{-1}$ sera la plus petite boule de $M_{(n,n)}$ contenant une matrice singulière.

Nous en déduisons immédiatement :

Théorème II,6

La distance -au sens d'une norme S_{pq}- d'une matrice A_o à l'ensemble A^* des matrices singulières est :

$$d(A_o, A^*) = \operatorname*{Inf}_{A \in A^*} S_{pq}(A - A_o) = \left[S_{pq}\,(A_o^{-1})\right]^{-1}$$

Ce résultat a été déjà obtenu par RIGAL et FRANCK[10], à partir de toute autre considération.

II.2 De même :

Théorème II.7

Pour que le lieu A_{A_o} des matrices carrées $\tilde{A}$ indiscernables de A_o au sens d'une norme H_p, $H_p(\tilde{A} - A_o) \leq \alpha$ contienne une matrice singulière, il faut et il suffit que

$$S_{H_q H_p}(A_o^{-1}) \geq \alpha^{-1}$$

La démonstration de ce théorème est analogue à celle du théorème II.5.

On en déduit un théorème équivalent au théorème II.6

Théorème II.8

La distance -au sens d'une norme H_p- d'une matrice A_o à l'ensemble A^* des matrices singulières est :

$$d(A_o, A^*) = \inf_{A \in A^*} H_p(A_o - A) = \left[S_{H_q H_p}(A_o^{-1})\right]^{-1}$$

III - QUELQUES CRITERES DE COMPATIBILITE

Les lemmes que nous venons de démontrer vont nous permettre d'établir quelques conditions nécessaires et suffisantes de compatibilité d'une solution calculée $\tilde{x}$ avec les données d'un système linéaire (A_o, b_o).

Nous avons vu dans l'introduction de cet ouvrage que les fonctions N_1 et N_2 caractérisaient le lieu des matrices indiscernables de A_o et celui des vecteurs indiscernables de b_o. Nous allons d'abord étudier le cas où la norme N_1 de l'application A est associée à deux normes vectorielles p et q.

$$N_1 (A) = S_{pq} (A)$$

III.1 - Critère de compatibilité dans le cas d'une norme S_{pq}.

Le lieu A_{A_o} des matrices $\tilde{A}$, indiscernables de A_o est caractérisé par :

$$A_{A_o} = \{ \tilde{A} \in M_{(m,n)} / \tilde{A} = A_o + A \; ; \quad S_{pq} (A) \leq \alpha \}$$

III.1.1 - Premier cas

Le lieu B_{b_o} des vecteurs $\tilde{b}$, indiscernables de b_o est défini par la norme p caractérisant la norme S_{pq} $[N_2(x) = p(x)]$

$$B_{b_o} = \{ \tilde{b} \in C^m / \tilde{b} = b_o + b \; ; \quad p(b) \leq \beta \}$$

Nous devons donc rechercher ici si, pour $\tilde{x}$ donné, le système suivant

$$(I) \quad \left| \begin{array}{l} (A_o + A)\,\tilde{x} = b_o + b \\ S_{pq}(A) \leq \alpha \\ p(b) \leq \beta \end{array} \right. \qquad \left[\begin{array}{l} A_o,\ b_o,\ S_{pq},\ p,\ \alpha,\ \beta, \\ \text{données du problème} \end{array} \right]$$

admet -au moins- une solution en A et b

D'après le lemme 2 (théorème II.2), l'ensemble de vecteurs y de C^m tels que

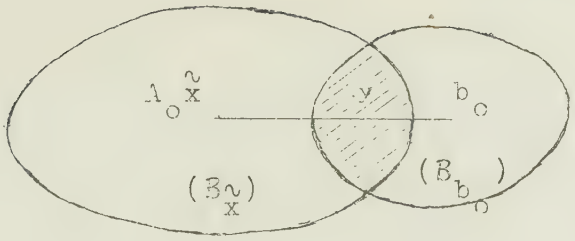
$$y = (A_o + A)\,\tilde{x}, \qquad S_{pq}(A) \leq \alpha$$

est la boule $B_{\tilde{x}}$ correspondant à la norme p, de centre $A_o\tilde{x}$ et de rayon $\alpha q(\tilde{x})$

$$\boxed{B_{\tilde{x}} = \{y \in C^m \ / \ p(y - A_o\tilde{x}) \leq \alpha q(\tilde{x})\}}$$

Ainsi pour que $\tilde{x}$ soit compatible avec les données il faut et il suffit que les deux boules $B_{\tilde{x}}$ et B_{b_o}, correspondant à la même norme p, (de centres $A_o\tilde{x}$ et b_o, de rayons $\alpha q(\tilde{x})$ et β) aient une intersection non vide, donc que la distance de leurs centres soit inférieure ou égale à la somme de leurs rayons, ce qui s'écrit

$$\boxed{d(A_o\tilde{x}, b_o) = p(A_o\tilde{x} - b_o) \leq \alpha q(x) + \beta}$$



ceci nous conduit à la condition, nécessaire et suffisante de compatibilité suivante :

$\tilde{x}$ compatible avec les données de (A_o, b_o)

$$\exists A, b : \begin{cases} (A_o + A)\,\tilde{x} = b_o + b \\ S_{pq}(A) \leq \alpha \\ p(b) \leq \beta \end{cases} \iff p(A_o \tilde{x} - b_o) \leq \alpha q(\tilde{x}) + \beta$$

III.1.2-Matrice et vecteur typiques

Lorsque $\tilde{x}$ est compatible avec les données, il peut être intéressant de déterminer *une* matrice A et *un* vecteur b qui correspondent au système (I).

Or nous savons que tout vecteur y appartenant à l'intersection des deux boules $\mathcal{B}_{\tilde{x}}$ et $\mathcal{B}_{b_o}$ détermine un tel couple ; A(y) et b(y) étant alors déterminés par

$$A(y) = \frac{y - A_o \tilde{x}}{\tilde{z}^H \tilde{x}} \cdot \tilde{z}^H$$

$$b(y) = b_o - y$$

A(y) étant d'ailleurs la matrice minimale (I.4) telle que

$$S_{pq}\ A(y) = \inf_{[A : y=(A_o+A)\tilde{x}]} S_{pq}(A) = \frac{p(y-A_o\tilde{x})}{q(\tilde{x})}$$

Or, si nous prenons le vecteur y_1 tel que

$$y_1 = \lambda A_o \tilde{x} + (1-\lambda)\, b_o$$

$$\text{avec} \qquad \lambda = \frac{\beta}{\alpha q(\tilde{x}) + \beta} \qquad (o \leq \lambda \leq 1)$$

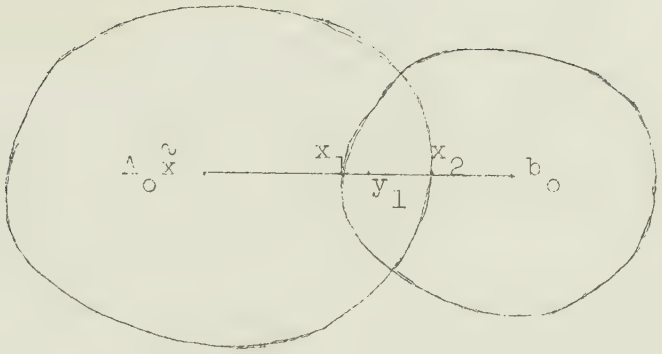
nous avons:

$$\left.\begin{aligned} p(y_1 - A_o\tilde{x}) &= (1-\lambda)\, p(A_o\tilde{x} - b_o) \leq (1-\lambda)\left[\alpha q(\tilde{x}) + \beta\right] = \alpha q(\tilde{x}) \\ p(y_1 - A_o) &= \lambda\, p(A_o\tilde{x} - b_o) \leq \lambda\left[\alpha q(\tilde{x}) + \beta\right] = \beta \end{aligned}\right\} \Longrightarrow y_1 \in \mathcal{B}_{\tilde{x}} \cap \mathcal{B}_{b_o}$$

et

$$\frac{p(y_1 - A_o\tilde{x})}{p(y_1 - b_o)} = \frac{\lambda - 1}{\lambda} = \frac{\alpha q(\tilde{x})}{\beta}$$

Ainsi les distances de y_1 aux centres $A_o\tilde{x}$ et b_o sont <u>proportionnelles aux rayons</u> $\alpha q(\tilde{x})$ et β.



y_1 sera d'autant plus proche de $A_o\tilde{x}$ (resp. de b_o) que α (resp. β) sera petit

Ceci nous a conduit à prendre la matrice A_1 et le vecteur b_1, correspondant à y_1 comme couple de pondération pour le système (I) :

$$y_1 = \frac{\beta}{\alpha q(\tilde{x})+\beta} A_o\tilde{x} + \frac{\alpha q(\tilde{x})}{\alpha q(\tilde{x})+\beta} b_o$$

$$A_1 = -\frac{\alpha q(\tilde{x})}{\left[\alpha q(\tilde{x})+\beta\right] z^H . \tilde{x}} \tilde{r} . z^H$$

$$b_1 = \frac{\beta}{\alpha q(\tilde{x}) + \beta} \tilde{r}$$

$(\tilde{r} = A_o\tilde{x} - b_o)$

III.1.3 Cas général

Lorsque les boules $\mathcal{B}_{\tilde{x}}$ et $\mathcal{B}_{b_o}$ ne sont pas homothétiques ($N_2 \neq p$), aucune condition nécessaire et suffisante ne peut être posée ; nous ne pouvons que recourir à l'équivalence des normes N_2 et p $\left[\forall x \in C^n \ : \ mp(x) \leq N_2(x) \leq Mp(x)\right]$ c'est-à-dire à des théorèmes d'inclusion de $\mathcal{B}_{b_o}$ et des boules $\mathcal{B}_1$ et $\mathcal{B}_2$, de même centre et de rayons respectifs $\beta/_m$ et $\beta/_M$, relativement à p.

$$\mathcal{B}_1 \subset \mathcal{B}_{b_o} \subset \mathcal{B}_2$$

On en déduit une condition nécessaire et une condition suffisante :

$\tilde{x}$ compatible avec les données de (A_o, b_o)

$$\exists A,b \ : \ \begin{cases} (A_o+A)\tilde{x} = b_o+b \\ S_{pq}(A) \leq \alpha \\ N_2(b) \leq \beta \end{cases} \quad \begin{array}{l} \Longrightarrow \ p(A_o\tilde{x}-b_o) \leq \alpha q(\tilde{x}) + \beta/_M \\ \\ \Longleftarrow \ p(A_o\tilde{x}-b_o) \leq \alpha\, q(\tilde{x}) + \beta/_m \end{array}$$

III.2 - Critère de compatibilité dans le cas des normes de Holder

Nous pouvons faire une étude analogue à la précédente dans le cas où le lieu A_{A_o} des matrices indiscernables de A_o est caractérisé par une norme de Holder d'ordre p sur $M_{(m,n)}$

$$A_{A_o} = \{\tilde{A} \in M_{(m,n)} \ / \ \tilde{A} = A_o + \Delta \ ; \ H_p(\Delta) \leq \alpha\}$$

III.2.1 - Lorsque le lieu B_{b_o} est caractérisé lui aussi par une norme de Holder d'ordre p sur C^m :

$$B_{b_o} = \{\tilde{b} \in C^m \ / \ \tilde{b} = b_o + b \ ; \ H_p(b) \leq \beta\}$$

nous avons par analogie avec ce qui précède (§ III.1.1)

$\tilde{x}$ compatible avec les données de (A_o, b_o)

$$\exists \Delta, b : \begin{cases} (A_o + \Delta)\,\tilde{x} = b_o + b \\ H_p(\Delta) \leq \alpha \\ H_p(b) \leq \beta \end{cases} \iff H_p(A_o\tilde{x} - b_o) \leq \alpha H_q(\tilde{x}) + \beta$$

III.2.2. - Matrice et vecteur typiques

Le coupe (A_1, b_1) sera alors défini par :

$$A_1 = \frac{-\alpha H_q(\tilde{x})}{[\alpha H_q(\tilde{x}) + \beta] H_p(\tilde{x})} \; \tilde{r}.\tilde{z}^H$$

$$b_1 = \frac{\beta}{\alpha H_q(\tilde{x}) + \beta} \; \tilde{r}$$

Il est à remarquer ici que le vecteur $\tilde{z}$ est entièrement déterminer par :

$$\tilde{z} = \frac{1}{[H_p(\tilde{x})]^{\frac{p}{q}}} \begin{vmatrix} \frac{|x_1|^p}{\tilde{x}_1} \\ \vdots \\ \frac{|\tilde{x}_n|^p}{\tilde{x}_n} \end{vmatrix}$$

et que

$$| \tilde{z}^H . \tilde{x} | = H_p(\tilde{x})$$

(c.f. chapitre I page 15)

III.2.3 - Dans le cas où la norme N_2 est quelconque nous avons avec les bornes m et M d'équivalence de N_2 et H_p

$\tilde{x}$ compatible avec les données de (A_o , b_o)

$$\exists A, b : \begin{cases} (A_o+A)\tilde{x} = b_o+b \\ H_p(A) \leq \alpha \\ N_2(b) \leq \beta \end{cases} \Longrightarrow H_p(A_o\tilde{x}-b_o) \leq \alpha H_q(\tilde{x})+\beta/m$$

$$\Longleftarrow H_p(A_o\tilde{x}-b_o) \leq \alpha H_q(\tilde{x})+\beta/M$$

III.3 - Critère de compatibilité dans le cas des normes vectorielles

La notion de normes vectorielles nous permet d'établir des conditions de compatibilité lorsque, la matrice A_o et le vecteur b_o se décomposant par "blocs", nous connaissons des bornes d'incertitude sur ces blocs. Les bornes seront données par la "matrice de majoration d'erreur" S (jouant le rôle analogue de α) et par le vecteur de majoration d'erreur B (jouant le rôle de β)

Nous considérons dans ce qui suit lesnormes vectorielles p d'ordre S, et q d'ordre r ; les espaces C^n et C^m seront décomposés comme dans le lemme 3 (théorème II.3) et nous emploierons les mêmes notations :

Décomposons A_o et b_o :

$$A_o = \begin{array}{|c|c|c|} \hline A^o_{11} & & A^o_{1r} \\ \hline & & \\ & & \\ \hline A^o_{s1} & & A^o_{sr} \\ \hline \end{array}$$

$$b_o = \left| \begin{array}{c} b^o_1 \\ \vdots \\ b^o_s \end{array} \right| \qquad \text{avec} \quad b^o_i \in C^{\beta i}$$

Le lieu des matrices $\tilde{A}$ indescernables de A_o est déterminé par :

$$A_{A_o} = \{ \tilde{A} \in M_{(m,n)} \; / \; \tilde{A} = A_o + A \; ; \quad M(A) \leq S \}$$

avec $M(A) = S_{p_i \, q_j} (A_{ij})$ et $S = (\alpha_{ij})$ $\quad i=1\text{-----}s \quad j=1\text{-----}r$

III.3.1 - Le lieu des vecteurs $\tilde{b}$, indiscernables de b_o est déterminé par :

$$\mathcal{B}_{b_o} = \{ \tilde{b} \in C^m \; / \; \tilde{b} = b_o + b \; ; \quad p(b) \leq B \}$$

avec $B \in R^s_+$; $B = \left| \begin{array}{c} \beta_1 \\ \vdots \\ \beta_s \end{array} \right|$

que nous décomposons en s boules $K_i \subset C^{\beta_i}$

$$K_i = \{\tilde{b}_i \in C^{\beta_i} \;/\; \tilde{b}_i = b_i^o + b_i \;;\; p_i(b_i) \leq \beta_i\} \qquad (i=1\text{-----}s)$$

La boule $\mathcal{B}_{\tilde{X}}$ devient :

$$\boxed{\mathcal{B}_{\tilde{X}} = \{Y \in C^m \;/\; p(Y - A_o \tilde{X}) \leq S.q\,(\tilde{X})\}}$$

que nous décomposons en s boules $H_i \subset C^{\beta_i}$

$$H_i = \{y_i \in C^{\beta_i} \;/\; p_i(y_i - [A_o \tilde{X}]_i) \leq \sum_{k=1}^{r} \alpha_{ik}\, q_k\,(\tilde{x}_k)\}$$

$$(i = 1\text{-----}\ s)$$

en notant $[A_o \tilde{X}]_i$ le vecteur de C^{β_i} tel que :

$$[A_o \tilde{X}]_i = \sum_{k=1}^{r} A_{ik}^o \cdot \tilde{x}_k$$

Les p_i étant des normes sur C^{β_i}, nous avons :

$$\mathcal{B}_{\tilde{X}} \cap \mathcal{B}_{b_o} \neq \phi \iff H_i \cap K_i \neq \phi \qquad (i=1\text{----}s)$$

Or sur chaque C^{β_i}, pour que l'intersection de H_i (de rayon $\sum_{k=1}^{r} \alpha_{ik}\, q_k(\tilde{x}_k)$ et de centre $[A_o \tilde{X}]_i$) et de K_i (de rayon β_i et de centre b_i^o) soit non vide, il faut et il suffit que :

$$p_i([A_o \tilde{X}]_i - b_i^o) \leq \sum_{k=1}^{r} \alpha_{ik}\, q_k(\tilde{x}_k) + \beta_i$$

Ceci étant valable pour toute intersection non vide $H_i \cap K_i$ (i=1----s), nous avons les conditions nécessaires et suffisantes de compatibilité :

$\tilde{x}$ compatible avec les données de (A_o, b_o)

$$\exists A, b \;:\; \begin{cases} M(A) \leqslant S \\ \quad \text{(coef. à coef.)} \\ p(b) \leqslant B \\ \quad \text{(comp. à comp.)} \end{cases} \iff p(A_o \tilde{X} - b_o) \leqslant S.q(\tilde{X}) + B$$

(inégalité vectorielle comp. à comp.)

II.3.2. Matrice et vecteur typiques

Le couple de pondération sera choisi comme précédemment, la matrice A_1 se décomposant en "matrice-bloc" $A^1_{ij} \in M_{(\alpha_i, \beta_j)}$ et le vecteur b_1 en $b^1_i \in C^{\beta_i}$

Nous avons avec les vecteurs $\tilde{Z}$ et $\tilde{Z}_{ij}$ définis comme au § I.2

$$A^1_{ij} = \frac{\sum_{k=1}^{r} \alpha_{ik}\, q_k(\tilde{x}_k)}{\left[\sum_{k=1}^{r} \alpha_{ik}\, q(\tilde{x}_k) + \beta_i\right] \tilde{Z}^H \tilde{X}} \; \tilde{r}_i \cdot \tilde{Z}^H_{ij} \qquad \begin{matrix} (i=1----s) \\ (j=1----r) \end{matrix}$$

et

$$b_i^1 = \frac{\sum_{k=1}^{r} \alpha_{ik}\, q_k\,(\tilde{x}_k)}{\sum_{k=1}^{r} \alpha_{ik}\, q_k\,(\tilde{x}_k) + \beta_i}\ \tilde{r}_i \qquad (i=1---s)$$

$$\text{avec} \quad \tilde{r} = \Lambda_o \tilde{X} - b_o = \begin{vmatrix} \tilde{r}_1 = [\Lambda_o \tilde{X}]_1 - b_1^o \in C^{\beta_1} \\ \vdots \\ \tilde{r}_s = [\Lambda_o \tilde{X}]_s - b_s^o \in C^{\beta_s} \end{vmatrix}$$

II.3.3 Nous pouvons choisir une norme vectorielle N, quelconque, d'ordre r (de composante N_i) pour déterminer $\mathcal{B}_{b_o}$.

$$\boxed{\mathcal{B}_{b_o} = \{\tilde{b} \ / \tilde{b} = b_o + b \ ; \quad N(b) \leq B\}}$$

Nous savons que chaque N_i est une norme sur C^{β_i}, il existe donc des constantes m_i et M_i telles que :

$$\forall x_i \in C^{\beta_i} : \quad m_i\, p_i(x_i) \leq N_i(x_i) \leq M_i\, p_i(x_i)$$

$$(i=1----s)$$

En considérant sur C^{β_i}, les boules de rayon β_i/m_i et β_i/M_i et leur inclusion avec $\mathcal{B}_{b_o}$ comme nous l'avons fait au §II.1.3, nous avons les conditions suivantes de compatibilité :

$\tilde{x}$ compatible avec les données de (A_o, b_o)

$$\exists\, A, b \;:\; \begin{cases} (A_o+A)\tilde{X} = b_o+b \\ M(A) \leq S \quad \text{(coef. à coef.)} \\ N(b) \leq B \quad \text{(comp. à comp.)} \end{cases} \Longrightarrow p(A_o\tilde{X}-b_o) \leq S\, q(\tilde{X}) + B_n \quad \text{(comp. à comp.)}$$

$$\Longleftarrow p(A_o\tilde{X}-b_o) \leq S\, q(\tilde{X}) + B_M$$

avec $B_m \in R_+^s$; $B_m = \begin{vmatrix} \beta_1/m_1 \\ \vdots \\ \beta_s/m_s \end{vmatrix}$ et $B_M \in R_+^s$; $B_M = \begin{vmatrix} \beta_1/M_1 \\ \vdots \\ \beta_s/M_s \end{vmatrix}$

II.3.4. Cas particuliers

1. Lorsque les normes vectorielles p et q se ramènent à des applications de C^m et C^n dans R_+ [r=s=1], nous retrouvons les résultats précédents.

2. En prenant r=s=n, nous obtenons les résultats d'OETTLI et PRAGER [11] qui considèrent une majoration d'erreur sur chaque composante de A_o et b_o.

Les normes vectorielles p et q deviennent :

$$p(x) = q(x) = \begin{vmatrix} |x_1| \\ \vdots \\ |x_n| \end{vmatrix}$$

et $M(A) = |A| = (|a_{ij}|)$

Les conditions de compatibilité sont alors :

$$(A_o + \mathbf{A})\,\tilde{x} = b_o + b$$

$$|a_{ij}| \leq \alpha_{ij} \quad (i,j=1---n) \iff |\tilde{r}_i| \leq \sum_{i=1}^{n} \alpha_{ij} |\tilde{x}_j| + \beta_i$$

$$|b_i| \leq \beta_i \quad (i=1---n) \qquad\qquad (i=1-----n)$$

et le couple de pondération

$$A_1 = (a^1_{ij}) \; ; \quad a^1_{ij} = - \frac{\tilde{r}_i \, \alpha_{ij} \, \dfrac{|\tilde{x}_i|}{x_i}}{\sum_{j=1}^{n} \alpha_{ij} |\tilde{x}_i| + \beta_i}$$

$$b_1 = (b^1_i) \; ; \quad b^1_i = \frac{\beta_i}{\Sigma \alpha_{ij} |\tilde{x}_i| + \beta_i} \, \tilde{r}_i$$

Tableau récapitulatif dans le cas des normes les plus employées.

Lieux d'incertitude de A_o, b_o	Conditions de compatibilité
$S_{\infty\infty}$: $\underset{i}{\mathrm{Max}}(\sum_j \lvert a_{ij}\rvert) \leq \alpha$ H_∞ : $\underset{i}{\mathrm{Max}} \lvert b_i\rvert \leq \beta$	$\underset{i}{\mathrm{Max}} \lvert\tilde{r}_i\rvert \leq \alpha . \underset{i}{\mathrm{Max}} \lvert\tilde{x}_i\rvert + \beta$
S_{11} : $\underset{j}{\mathrm{Max}}(\sum_i \lvert a_{ij}\rvert) \leq \alpha$ H_1 : $\sum_i \lvert b_i\rvert \leq \beta$	$\sum_i \lvert\tilde{r}_i\rvert \leq \alpha \sum_i \lvert\tilde{x}_i\rvert + \beta$
S_{22} : $\lvert\rho(A^H A)\rvert^{\frac{1}{2}} \leq \alpha$ H_2 : $(\sum_i \lvert b_i\rvert^2)^{\frac{1}{2}} \leq \beta$	$(\sum_i \lvert\tilde{r}_i\rvert^2)^{\frac{1}{2}} \leq \alpha . (\sum_i \lvert\tilde{x}_i\rvert^2)^{\frac{1}{2}} + \beta$
$S_{\infty 2}$: $\underset{i}{\mathrm{Max}}(\sum_k \lvert a_{ik}\rvert^2)^{\frac{1}{2}} \leq \alpha$ H_∞ : $\underset{i}{\mathrm{Max}} \lvert b_i\rvert \leq \beta$	$\underset{i}{\mathrm{Max}} \lvert\tilde{r}_i\rvert \leq \alpha (\sum_i \lvert\tilde{x}_i\rvert^2)^{\frac{1}{2}} + \beta$
H_∞ : $\underset{ij}{\mathrm{Max}} \lvert a_{ij}\rvert \leq \alpha$ H_∞ : $\underset{i}{\mathrm{Max}} \lvert b_i\rvert \leq \beta$	$\underset{i}{\mathrm{Max}} \lvert\tilde{r}_i\rvert \leq \alpha \sum_i \lvert\tilde{x}_i\rvert + \beta$
H_1 : $\sum_{ij} \lvert a_{ij}\rvert \leq \alpha$ H_1 : $\sum_i \lvert b_i\rvert \leq \beta$	$\sum_i \lvert\tilde{r}_i\rvert \leq \alpha \underset{i}{\mathrm{Max}} \lvert\tilde{x}_i\rvert + \beta$
H_2 : $(\sum_{ij} \lvert a_{ij}\rvert^2)^{\frac{1}{2}} \leq \alpha$ H_2 : $(\sum_i \lvert b_i\rvert^2)^{\frac{1}{2}} \leq \beta$	$(\sum_i \lvert\tilde{r}_i\rvert^2)^{\frac{1}{2}} \leq \alpha (\sum_i \lvert\tilde{x}_i\rvert^2)^{\frac{1}{2}} + \beta$
Normes vectorielles $\lvert a_{ij}\rvert \leq \alpha_{ij}$ $(i,j=1---n)$ $\lvert b_i\rvert \leq \beta_i$ $(i=1---n)$	$\tilde{r}_i \leq \sum_j \alpha_{ij} \lvert\tilde{x}_j\rvert + \beta_i$ $(i=1---n)$

IV - APPLICATION AUX RACINES D'UN POLYNOME

Le critère de compatibilité peut s'appliquer à la recherche des racines d'un polynôme P_o, de coefficients a_o ------ a_n

$$P_o(x) = a_n x^n + \text{------} + a_o$$

Nous ferons une étude matricielle du problème en posant :

$$A = \begin{vmatrix} a_o \\ \vdots \\ a_n \end{vmatrix} \qquad X = \begin{vmatrix} 1 \\ x \\ \vdots \\ x^n \end{vmatrix}$$

$$P_o(x) = o \iff A^t . X = o$$

Soit **p** un polynôme de coefficients $a_o+\delta a_o$ ------- $a_n+\delta a_n$

et :

$$\Delta A = \begin{vmatrix} \delta a_o \\ \vdots \\ \delta a_n \end{vmatrix}$$

On dira que le polynôme P est indiscernable-au sens d'une norme q-de P_o si :

$$q(\Delta A) \leq \alpha$$

et $\tilde{x}$ sera accepté comme racine du polynôme P_o si $\tilde{x}$ est racine d'un polynôme indiscernable de P_o, c'est-à-dire matriciellement, s'il existe un vecteur ΔA tel que :

$$\left| \begin{array}{l} (A_o + \Delta A)^t \tilde{X} = o \\ q(\Delta A) \leq \alpha \\ \\ \text{avec} \quad \tilde{X} = \begin{vmatrix} 1 \\ \tilde{x} \\ \vdots \\ \tilde{x}^n \end{vmatrix} \end{array} \right.$$

En remarquant que toute norme q peut s'écrire pour $A \in M_{(n+1,1)}$

$$q(A) = \sup_{X \in C^{n+1}} \frac{|A^t X|}{q^{\#}(X)} = S_{H_1} q^{\#}(A) \qquad \text{II.2)}$$

et en tenant compte du lemme 2 (théorème, nous pouvons donné une condition nécessaire et suffisante de compatibilité de $\tilde{x}$:

$\tilde{x}$ compatible avec les données de P_o

$$\exists \Delta A : \left| \begin{array}{l} (A_o + \Delta A)\tilde{X} = o \\ q(\Delta A) \leq \alpha \end{array} \right. \iff |A_o^t \tilde{X}| \leq \alpha\, q^{\#}(\tilde{X})$$

Le vecteur typique ΔA, sera :

$$\Delta A_1 = \frac{\Psi_o(\tilde{x})}{\tilde{Z}^H \tilde{X}} \cdot \tilde{Z} \qquad \text{en remarquant qu'ici } \tilde{r} = P_o(\tilde{x}) \in C$$

Lorsque q est une norme de Holder d'ordre p : H_p ,

$$\tilde{Z} = \frac{1}{|H_p(\tilde{X})|^{\frac{p}{q}}} \begin{vmatrix} \frac{|x_1|^p}{x_1} \\ \vdots \\ \frac{|x_n|^p}{x_n} \end{vmatrix}$$

et nous obtenons :

$$\Delta A_1 = \frac{P_o(\tilde{x})}{[H_p(\tilde{X})]^p} \cdot \begin{vmatrix} 1 \\ \frac{|\tilde{x}|^p}{\tilde{x}} \\ \vdots \\ \frac{|\tilde{x}^n|^p}{\tilde{x}^n} \end{vmatrix}$$

Dans le cas où nous avons K racines calculées ($K \leq n$) $\tilde{x}_1$ --- $\tilde{x}_k$, nous obtenons les K conditions :

$$|A_o^t \tilde{X}_i| \leq \alpha\, q^*(\tilde{X}_i) \qquad (i=1\text{----}K)$$

mais ici l'existence d'un polynôme unique indiscernable n'est pas assuré. Une étude dans ce cas a été faite par NGUYEN HUU VINH [12] qui a pu résoudre le problème en faisant une généralisation d'une semi-norme duale (q = norme euclidienne).

CHAPITRE III

ETUDE PRATIQUE

Les méthodes de résolution d'un système linéaire peuvent se diviser en deux grandes catégories : les méthodes directes et les méthodes indirectes. Nous allons étudier ici le rôle de"l'analyse à postériori" appliquée à chaque méthode.

1- Méthodes Directes

Elles consistent à transformer le système (A_o, b_o) en un système plus simple et plus facile à résoudre, admettant même solution.

$$A_o\, x = b_o \longrightarrow T(A_o)x = T(b_o)$$

C'est le cas des méthodes de Gauss et des méthodes d'orthogonalisation (processus finis).

Pour"l'analyse à postériori", il suffira de vérifier par exemple que

$$p(\tilde{r}) \leq \alpha\, q(\tilde{x}) + \beta$$

dans le cas où les données sont A_o, b_o, Spq, p, α, β .

Notons que ce test est facilement calculable si l'on prend pour p et q des normes de Holder.

Dans le cas où

$$p(\tilde{r}) > \alpha\, q(\tilde{x}) + \beta$$

nous devons rejeter $\tilde{x}$ comme solution (ou nous astreindre à augmenter les bornes α et β si $p(\tilde{r}) \neq\neq \alpha\, q(\tilde{x}) + \beta$).

De telles méthodes ne sont utilisables en général que pour des systèmes d'ordre inférieur à 50 . En effet la transformation de (A_o, b_o) nécessite un nombre assez considérable d'opérations, en particuler des multiplications ; il s'en suit une accumulation importante des erreurs.

2- Méthodes indirectes

Pour de grands systèmes, fréquents dans la pratique (systèmesd'ordre 1000 obtenus pour la résolution d'équations aux différences données par des équations différentielles),

on emploie des méthodes indirectes, qui consistent à "s'approcher " de plus en plus de la solution du système proposé.

On part d'un vecteur arbitraire x_1 ; on forme le résidu $r_1 = A_o\, x_1 - b_o$; on cherche un moyen de modifier x_1 en x_2(donc r_1 en r_2) et d'une manière générale x_m en x_{m+1}(donc r_m en r_{m+1})de telle façon que

$$p(r_{m+1}) \quad < p(r_m)$$

p étant une norme quelconque sur C^n.

Si le moyen de passer de x_m à x_{m+1} n'altère qu'une seule composante de x_m, on dira que l'on fait une relaxation sinon on parlera d'itération.

Dans le cas où l'on emploie une analyse directe nous arrêterons le processus lorsque $p(r_m)$ sera suffisamment petit on pourra dire alors que x_m est suffisamment proche de la solution théorique.

En effet une condition nécessaire et suffisante pour que la suite $\{x_m\}$ tende vers x est que la suite $\{r_m\}$ tende vers o.

Preuve

Nous supposerons évidemment que A_o n'est pas singulière.

Soit : $A_o\, x = b_o$

$$r_m = A_o\, x_m - b_o = A_o\, (x_m - x)$$

et

1) $\quad p(r_m) = p\,[A_o\, (x_m - x)] \leq Spq(A_o)\; q(x_m - x)$

d'où : Si $\quad x_m \longrightarrow x$

$$q\,(x_m - x) \longrightarrow o \Longrightarrow p(r_m) \longrightarrow o$$

donc $\quad r_m \longrightarrow o.$

$$2)\quad \frac{p(r_m)}{q(x_m-x)} = \frac{p\,[A_o(x_m-x)]}{q(x_m-x)} \geqslant \underset{\substack{x\in C^n \\ (x_m-x)\neq o}}{\text{Inf.}} \frac{p\,[A_o(x_m-x)]}{q(x_m-x)}$$

$$= \frac{1}{\underset{\substack{x\in C^n \\ (x_m-x)\neq o}}{\text{Sup}} \dfrac{q[(x_m-x)]}{p[A_o(x_m-x)]}} = \frac{1}{\text{Spq}\ (A_o^{-1})}$$

d'où

$$q(x_m - x) \leqslant \text{Sqp}\ (A_o^{-1})\ p\ (r_m)$$

et si $r_m \longrightarrow o$: $p(r_m) \longrightarrow o \Longrightarrow q(x_m-x) \longrightarrow o$

donc $x_m \longrightarrow x$.

Nous voyons que dans ces méthodes, au contraire des méthodes directes, *c'est au cours des calculs* que nous avons un moyen de déterminer la validité d'une solution.

Pour l'analyse à postériori, dès que nous avons trouvé un vecteur x_m vérifiant le critère de compatibilité, nous arrêterons le processus.Il serainutile d'aller plus loin ; le fait de continuer les calculs ne présente plus aucun intérêt. En effet, notre but étant de trouver la solution exacte *d'un système quelconque* appartenant aux lieux d'incertitudes des données, tout vecteur x_{m+n} possèdera les mêmes propriétés que x_m et ne nous apportera aucun renseignement supplémentaire.

Les données étant incertaines, il n'est pas plus logique de résoudre (A_o, b_o) que tout autre système du lieu ; le système (A_o, b_o) ne joue ici qu'un rôle de "repère".

Le critère $p(r_m) \leq \alpha q(\tilde{x}_m) + \beta$, étant mathématiquement plus rigoureux que le critère $p(r_m) \leq \varepsilon$ permet de gagner un nombre parfois appréciable d'itérations [ε - même s'il est déterminé par les incertitudes des données n'est en général qu'une majoration de l'erreur].

Nous allons ici exposer quelques méthodes indirectes de résolution et les appliquer à des systèmes donnés (à coefficients réels).

Nous donnerons dans chaque cas, un tableau de résultats qui nous permettra d'apprécier le gain obtenu sur le nombre d'itérations.

I - Méthode de Southwell (Relaxation)

I.1 - Principe

$$\text{Soit :} \quad x_m = \begin{vmatrix} \xi_1^m \\ \vdots \\ \xi_n^m \end{vmatrix} \quad ; \quad r_m = A_o x_m - b_o = \begin{vmatrix} \rho_1^m \\ \vdots \\ \rho_n^m \end{vmatrix}$$

Pour déterminer x_{m+1} à partir de x_m, on pose :

$$x_{m+1} = \begin{vmatrix} \xi_1^{m+1} \\ | \\ | \\ | \\ | \\ \xi_n^{m+1} \end{vmatrix}$$

avec $\xi_k^{m+1} = \xi_k^m$ (k = 1------- j-1, j+1----- n)

ξ_j^{m+1} est choisi de telle façon qu'il annule la i^{e} composante de r_m.

On a :

$$\xi_j^{m+1} = \frac{1}{a_{ij}} \left[b_i - \sum_{k=1}^{n} a_{ik} \xi_k^m \right] \qquad (a_{ij} \neq o)$$

ce qui donne

$$\xi_j^{m+1} = \xi_j^m - \frac{\rho_i^m}{a_{ij}}$$

On en déduit

$$\rho_k^{m+1} = \rho_k^m - a_{kj} \frac{\rho_i^m}{a_{ij}} \qquad (k = 1----- i-1, i+1--- n)$$

$$\rho_i^{m+1} = o$$

Pour la méthode de Southwell, on choisit l'indice tel que

$$|\rho_i^m| = \max_k |\rho_k^m|$$

et on prend $j = i$

Les formules deviennent :

$$\left.\begin{aligned} \xi_k^{m+1} &= \xi_k^m \\ \rho_k^{m+1} &= \rho_k^m - a_{ki} \frac{\rho_i^m}{a_{ii}} \end{aligned}\right| \quad (k=1\text{------}\ i-1\ ;\ i+1\text{----}\ n)$$

$$\xi_i^{m+1} = \xi_i^m - \frac{\rho_i^m}{a_{ii}}$$

$$\rho_i^{m+1} = 0$$

I.2 - Convergence [13]

Une condition suffisante de convergence de la méthode de Southwell est :

$$\max_j \sum_{\substack{i=1 \\ i\neq j}}^{n} \frac{|a_{ij}^o|}{|a_{jj}^o|} < \frac{1}{n}$$

Cette condition est bien trop rigoureuse. Mais dans le cas où A_o est symétrique définie positive et admet une diagonale composée des mêmes éléments, on peut affirmer que la méthode est toujours convergente quelque soit le vecteur x_1 de départ.

I.3 - Etude pratique

1. On utilisera dans l'étude suivante, les normes H_∞, H_∞ . Les conditions sur les lieux d'incertitude seront alors :

$$H_\infty(A) = \max_{i,j} |a_{ij}| \leqslant \alpha$$

$$H_\infty(b) = \max_{i} |b_i| \leqslant \beta$$

2. Procédure ALGOL

```
procedure RESOLTEST1(A) SECOND MEMBRE:(B) ORDRE:(N) VECTEUR
          SOLUTION:(X) INCERTITUDE SUR A:(ALFA) INCERTITUDE
          SUR B:(BETA) NORME R : (P)  NORME X : (Q);
entier N ; reel ALFA, BETA, P, Q; reel tableauA, B, X;
commentaireRésolution par une méthode itérative (SOUTHWELL) d'un
système linéaire dont les données (A,B) sont considérées comme
incertaines - (Norme de R : P=H∞   Norme de X : Q=H1);
debut procedure SOUTHWELL (A) SECOND MEMBRE:(B) ORDRE:(N)
                VECTEUR ITERE:(X) RESIDU:(R) NORME X:(Q)
                NORME R:(P) INDICE DU MAX  DE R:(K);
```

```
entier N, K; reel P, Q; reel tableau A, B, X, R;
commentaire Passage d'un vecteur itéré au suivant par la
méthode de SOUTHWELL;
début entier I, J,
        J := K;
        Q := Q - ABS (X [J] );
        X [J] := X [J] - R [J] / A [J,J] ;
        Q := Q + ABS (X [J] );
        P := 0.0;
        pour I := 1 pas 1 jusqua J-1, J+1 pas 1 jusqua N faire
        debut R [I] := R [I] - A [I,J] x R [J] / A [J,J] ;
        si ABS(R [I])≥P alors
        debut P := ABS(R [I]) ;
              K := I
        fin
        fin;
        R[I]:= 0.0
fin SOUTHWELL
        entier I, J, K; reel S, E; reel tableau R [1:N];
        P := 0.0 ;  Q := 0.0;
        pour I := 1 pas 1  jusqua N faire
        debut S := 0.0;
              pour J := 1 pas 1 jusqua N faire
              S := S + A [I,J] x X [J] ;
              R [I] := S - B [I] ;
              si ABS(R [I])≥ P alors
              debut  P := ABS (R [I]) ;
                     K := I
              fin;
              Q := Q + ABS (X [I])
        fin
```

```
BOUCLE :  SOUTHWELL (A, B, N, X, R, Q, P, K);
          E := ALFA x Q + BETA
          si P>E alors allera BOUCLE
     fin  RESOLTEST1;
```

3. Résultats numériques

Les tableaux suivants donnent pour différentes valeurs de α et β, le vecteur itéré compatible avec les données et le nombre I d'itérations.

La variable K indique le nombre d'itérations nécessaires lorsque l'on utilise comme critère $p(r_n) \leq \varepsilon (\varepsilon = \alpha = \beta)$. On souligne, à titre indicatif, les décimales inexactes par rapport à la solution théorique de (A_o, b_o).

(On emploiera le même modèle pour les autres méthodes).

Exemple 1

$$A_o = \begin{vmatrix} 4,33 & -1,12 & -1,08 & 1,14 \\ -1,12 & 4,33 & 0,24 & -1,22 \\ -1,08 & 0,24 & 7,21 & -3,22 \\ 1,14 & -1,22 & -3,22 & 5,43 \end{vmatrix} \quad b_o = \begin{vmatrix} 3,52 \\ 1,57 \\ 0,54 \\ -1,09 \end{vmatrix}$$

β \ α	10^{-8}	10^{-7}	10^{-6}	10^{-5}
10^{-8}	Précision non atteinte	I=41 1,046 248 91 0,562 782 88 0,111 002 69 -0,228 121 62	I=34 1,046 248 67 0,562 782 58 0,111 002 60 -0,228 121 77	I=27 1,046 250 70 0,562 781 57 0,111 001 84 -0,228 124 21
10^{-7}	I=41	I=40 1,046 248 91 0,562 782 88 0,111 002 67 -0,228 121 62 K=42	I=34 -	I=27 -
10^{-6}	I=36 1,046 248 91 0,562 782 82 0,111 002 60 -0,228 121 77	I=34 -	I=32 1,046 248 67 0,562 782 58 0,111 002 44 -0,228 122 13 K=36	I=27 -
10^{-5}	I=27 -	I=27 -	I=27 -	I=26 1,046 250 70 0,562 781 57 0,110 999 41 -0,228 124 21 K=27

Exemple 2

$$A_o = \begin{vmatrix} -13 & 2 & -1 & 3 \\ -6 & 19 & 4 & 5 \\ 1 & -5 & -37 & 7 \\ 5 & -9 & 8 & 23 \end{vmatrix} \quad b_o = \begin{vmatrix} -61 \\ -217 \\ 247 \\ 590 \end{vmatrix}$$

β \ α	10^{-8}	10^{-7}	10^{-6}	10^{-5}
10^{-8}	I=36 6,999 999 87 -13,999 999 73 - 1,000 000 01 19,000 000 00 K>40	I=32 6,999 999 81 -13,999 999 16 - 1,000 000 11 19,000 000 00	I=29 6,999 998 6 -13,999 999 04 - 1,000 000 11 19,000 000 00	I=24 6,999 991 4 -14,000 007 6 - 0,999 998 8 18,999 998 0
10^{-7}	I=36 -	I=32 - K=36	I=29 -	I=24 -
10^{-6}	I=33 6,999 999 81 -13,999 999 54 - 1,000 000 11 19,000 000 00	I=32 -	I=29 - K=34	I=24 -
10^{-5}	I=30 6,999 999 52 -13,999 999 04 - 1,000 000 11 19,000 000 00	I=30 -	I=29 -	I=24 - K=30

Exemple 3

$$A_o = \begin{vmatrix} 4 & -1 & -1 & 0 \\ -1 & 4 & 0 & -1 \\ -1 & 0 & 4 & -1 \\ 0 & -1 & -1 & 4 \end{vmatrix} \quad b_o = \begin{vmatrix} 1 \\ 0 \\ 0 \\ 0 \end{vmatrix}$$

β \ α	10^{-8}	10^{-7}	10^{-6}	10^{-5}
10^{-8}	Précision non atteinte	Précision non atteinte	I=41 0,291 666 626 0,083 333 25 0,083 333 25 0,041 666 626	I=35 0,291 666 03 0,083 333 01 0,083 333 01 0,041 666 03
10^{-7}	I=45 0,291 666 656 0,083 333 313 0,083 333 313 0,041 666 656	I=43 0,291 666 626 0,083 333 313 0,083 333 313 0,041 666 626 K=45	I=41 -	I=35 -
10^{-6}	I=37 0,291 666 50 0,083 333 01 0,083 333 01 0,041 666 50	I=37	I=37 K=37	I=33 0,291 666 03 0,083 332 0,083 332 0,041 666 03
10^{-5}	I=31 0,291 664 0,083 332 0,083 332 0,041 664	I=31 -	I=31 -	I=31 - K=31

II - Méthode de Gauss-Seidel

II.1 Principe

Cette méthode part du même principe que celle de Southwell ; seul le choix de i et j diffère.

Au $(m+1)^{e}$ pas, on choisit l'indice (qui annulera la i^{e} composante de r_m) égal au reste de la division de m+1 par n (ou égal à n si ce reste est nul).

On prend ici aussi l'indice j égal à i.

Ceci présente l'avantage de respecter l'ordre de numérotation et de pouvoir grouper aisément n pas en un seul.

Après n itérations, on dira qu'on est passé de X_m à X_{m+1} (en ayant utiliser n vecteurs intermédiaires)

La K^{e} composante ξ_k^{m+1} de X_{m+1} est alors :

$$\xi_k^{m+1} = \frac{1}{a_{kk}} \left[b_k - \sum_{e=1}^{i-1} a_{ke}\ \xi_e^{m+1} - \sum_{e=i+1}^{n} a_{ke}\ \xi_e^{m} \right]$$

$$(k = 1 \text{ -------- } n)$$

ou matriciellement

$$X_{m+1} = A_1^{-1} (b - A_2\ X_m)$$

$$\text{Avec } A_o = A_1 + A_2 \quad \begin{cases} A_1 = (a^1_{ij}) & \begin{cases} a^1_{ij} = a_{ij} & \text{pour } i \geqslant j \\ a^1_{ij} = o & \text{pour } i < j \end{cases} \\ A_2 = (a^2_{ij}) & \begin{cases} a^2_{ij} = a_{ij} & \text{pour } i < j \\ a^2_{ij} = o & \text{pour } i \leqslant j \end{cases} \end{cases}$$

II.2 - Convergence [14]

1. Pour que l'itération de Gauss-Seidel converge il faut et il suffit que le module de la plus grande valeur propre de $A_1^{-1} A_2$ soit strictement inférieur à l'unité.

2. On en déduit la condition suffisante suivante :

$$\operatorname*{Max}_i \sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{|a_{ik}|}{|a_{ii}|} < 1$$

II.3 - Etude pratique

1. Les normes utilisées sont ici $p=q=H_\infty$
Les lieux d'incertitudes seront :

$$S_{H_\infty H_\infty}(A) = \operatorname*{Max}_i \sum_j |a_{ij}| \leqslant \alpha$$

$$H_\infty(b) = \operatorname*{Max}_i |b_i| \leqslant \beta$$

2. Procédure ALGOL

```
procédure RESOLTEST (A) SECOND MEMBRE: (B) ORDRE: (N) VECTEUR
          SOLUTION: (X) INCERTITUDE SUR A: (ALFA) INCERTITUDE
          SUR B: (BETA) NORME R: (P) NORME X: (Q) ;

entier N; réel ALFA, BETA, P, Q, ; reel tableau A, B, X;
commentaire Résolution par une méthode itérative (Gauss-Seidel)
d'un système linéaire dont les données(A,B) sont considérées
comme incertaines (Norme du résidu: P=H∞, Norme du vecteur solution
Q=H∞) ;

début procédure GAUSSEIDEL (A) SECOND MEMBRE:(B) ORDRE:(N)
                VECTEUR ITERE:(X) NORME R:(P) NORME X:(Q) RESIDU:
                (R) INDICE DU COEFFICIENT NUL DU RESIDU: (K);
      entier N,K; reel P,Q; reel tableau A, B, X, R;
      commentaire Passage d'un vecteur itéré à l'itéré suivant par
      la méthode de Gauss-Seidel ;
      début  entier I; reel L;
             L := R[K]/A[K,K];
             X[K] := X[K]-L;
             si ABS(X[K])≥ Q  alors Q := ABS(X[K]);
             P := 0.0 ;
             pour I := 1 pas 1 jusqua K-1, K+1 pas 1 jusqua N faire
             début R[I]:= R[I]- A[I,K]x L;
                   si ABS(R[I])≥P  alors P := ABS(R[I])
             fin ;
             R[K]:= 0.0
      fin GAUSSEIDEL;

      entier I,J; reel K,S,E; reel tableau R[1:N];
```

```
        pour I := 1  pas 1  jusqua N faire
        début S := 0.0;
              pour J := 1 pas 1 jusqua N faire
              S := S + A[I,J]x X [J] ;
              R[I] := S - B[I]
        fin;
        Q := 0.0;
        pour I := 1  pas 1  jusqua N faire
        si ABS(X[I]) ≥Q  alors Q := ABS(X[I]);
        K := 0;
BOUCLE : K := K + 1;
        si K>N alors K := 1;
        GAUSSEIDEL (A, B, N, X, P, Q, R, K);
        E := ALFA x Q + BETA;
        si P>E  alors allera BOUCLE
fin RESOLTEST;
```

II. 3 Résultats numériques

Exemple 1

$$A_o = \begin{vmatrix} -6 & 5 & 1 \\ 2 & -4 & 1 \\ 2 & 3 & 1 \end{vmatrix} \qquad b_o = \begin{vmatrix} -4 \\ 3 \\ 10 \end{vmatrix}$$

α / β	10^{-8}	10^{-7}	10^{-6}	10^{-5}	10^{-4}
10^{-8}	Précision non atteinte	I=124 1,999 999 52 0,999 999 46 2,999 999 76	I=109 1,999 999 76 0,999 999 70 2,999 999 76	I=94 2,000 002 1 1,000 004 5 2,999 989 9	I=72 1,999 959 0,999 961 2,999 951
10^{-7}	I=135 1,999 999 52 0,999 999 52 3,000 000 00	I=124 - - K=135	I=109 -	I=94 -	I=72 -
10^{-6}	I=121 1,999 999 64 0,999 999 64 2,999 999 52	I=109 -	I=108 1,999 999 16 0,999 999 70 2,999 999 76 K=121	I=94 .	I=72 -
10^{-5}	i=96 1,999 998 4 0,999 998 2 2,999 998 8	I=96 -	I=96 -	I=93 1,999 997 3 1,000 004 5 2,999 989 1 K=96	I=72 -
10^{-4}	I=85 2,000 006 9 1,000 006 6 3,000 007 3	I=85 -	I=85 -	I=84 1,999 985 1,000 006 6 3,000 007 3	I=72 - K=81

Exemple 2

$$A_o = \begin{vmatrix} -13 & 2 & -1 & 3 \\ -6 & 19 & 4 & 5 \\ 1 & -5 & -37 & 7 \\ 5 & -9 & 8 & 23 \end{vmatrix} \qquad b_o = \begin{vmatrix} -61 \\ -217 \\ 247 \\ 590 \end{vmatrix}$$

β \ α	10^{-8}	10^{-7}	10^{-6}	10^{-5}
10^{-8}	I=47 6,999 999 04 -14,000 000 00 - 0,999 999 64 19,000 000 00 K=?	I=40 6,999 999 04 -14,000 000 00 - 0,999 999 58 19,000 000 00	I=35 6,999 993 5 -14,000 000 0 - 0,999 999 64 19,000 000 00	I=32 6,999 996 6 -14 000 002 8 - 1,000 001 5 19,000 000 0
10^{-7}	I=47 "	I=40 "	I=35 "	I=32 "
10^{-6}	I=43 6,999 999 04 -14,000 000 00 - 0,999 999 64 19,000 000 00	I=40 "	I=35 " K=44	I=32 "
10^{-5}	I=35 "	I=35 "	I=35 "	I=32 " K=36

Exemple 3

$$A_o = \begin{vmatrix} 4,33 & -1,12 & -1,08 & 1,14 \\ -1,12 & 4,33 & 0,24 & -1,22 \\ -1,08 & 0,24 & 7;21 & -3,22 \\ 1,14 & -1,22 & -3,22 & 5;43 \end{vmatrix} \qquad b_o = \begin{vmatrix} 3,52 \\ 1,57 \\ 0,54 \\ -1,09 \end{vmatrix}$$

β \ α	10^{-8}	10^{-7}	10^{-6}	10^{-5}	10^{-4}
10^{-8}	Précision non atteinte	I=55 1,046 249 03 0,562 782 94 0,111 002 70 -0,228 121 57	I=47 1,046 249 15 0,562 783 00 0,111 002 81 -0,228 121 37	I=39 1,046 250 70 0,562 783 89 0,111 004 93 -0,228 120 75	I=35 1,046 255 23 0,562 786 57 0,111 008 83 -0,228 109 67
10^{-7}	I=55 -	I=51 1,046 249 03 0,562 782 94 0,111 002 72 -0,228 121 57 K=55	I=47 -	I=39 -	I=35 -
10^{-6}	I=47 -	I=47 -	I=44 1,046 249 50 0,562 783 18 0,111 003 13 -0,228 121 37 K=47	I=39 -	I=35 -
10^{-5}	I=39 -	I=39 -	I=39 -	I=38 1,046 250 70 0,562 783 89 0,111 004 33 -0,228 118 43 K=39	I=35 -

Exemple 4

$$A_o = \begin{vmatrix} 4 & -1 & 0 & -1 & & & & & \\ -1 & 4 & -1 & 0 & -1 & & & & \\ 0 & -1 & 4 & 0 & 0 & -1 & & & \\ -1 & 0 & 0 & 4 & -1 & 0 & -1 & & \\ & -1 & 0 & -1 & 4 & -1 & 0 & -1 & \\ & & -1 & 0 & 0 & 4 & 0 & 0 & -1 \\ & & & -1 & 0 & -1 & 4 & -1 & 0 \\ & & & & -1 & 0 & -1 & 4 & -1 \\ & & & & & -1 & 0 & -1 & 4 \end{vmatrix}$$

$$b_o = \begin{vmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix}$$

β \ α	10^{-7}	10^{-6}	10^{-5}	10^{-4}
10^{-7}	I=176 -0,299 107 07 -0,098 214 22 -0,031 249 97 -0,098 214 22 -0,062 499 9 -0,026 785 6 -0,031 249 94 -0,026 785 65 -0,013 392 82 K=194	I=167 -0,299 107 01 -0,098 214 16 -0,031 249 94 -0,098 214 16 -0,062 499 88 -0,026 785 59 -0,031 249 88 -0,026 785 59 -0,013 392 79	I=140 -0,299 106 18 -0,098 213 32 -0,031 249 52 -0,098 213 32 -0,062 499 04 -0,026 784 76 -0,031 249 04 -0,026 784 76 -0,013 392 38	I=113 -0,299 099 50 -0,098 206 65 -0,031 246 18 -0,098 206 65 -0,062 492 37 -0,026 778 08 -0,031 242 37 -0,026 778 08 -0,013 389 04
10^{-6}	I=149 -0,299 106 65 -0,098 213 80 -0,031 249 76 -0,098 213 80 -0,062 499 52 -0,026 785 23 -0,031 249 52 -0,026 785 23 -0,013 392 61	I=149 - K=149	I=131 -0,299 105 22 -0,098 212 37 -0,031 249 04 -0,098 212 37 -0,062 498 09 -0,026 783 80 -0,031 248 09 -0,026 783 80 -0,013 391 90	I=104 -0,299 091 87 -0,098 199 02 -0,031 242 37 -0,098 199 02 -0,062 484 74 -0,026 770 45 -0,031 234 74 -0,026 770 45 -0,013 385 22
10^{-5}	I=122 -0,299 103 31 -0,098 210 46 -0,031 248 09 -0,098 210 46 -0,062 496 18 -0,026 781 89 -0,031 246 18 -0,026 781 89 -0,013 390 94	I=122 -	I=122 - K=122	I=104 -
10^{-4}	I=96 -0,299 076 61 -0,098 183 76 -0,031 234 74 -0,098 183 76 -0,062 469 48 -0,026 770 45 -0,031 219 48 -0,026 755 19 -0,013 377 59	I=96 -	I=96 -	I=86 -0,299 046 09 -0,098 153 24 -0,031 219 48 -0,098 153 24 -0,062 438 96 -0,026 724 67 -0,031 188 96 -0,026 724 67 -0,013 362 33 K=86

Exemple 5

$$A_o = \begin{vmatrix} 4 & -1 & -1 & 0 \\ -1 & 4 & 0 & -1 \\ -1 & 0 & 4 & -1 \\ 0 & -1 & -1 & 4 \end{vmatrix} \qquad b_o = \begin{vmatrix} 1 \\ 0 \\ 0 \\ 0 \end{vmatrix}$$

ε \ α	10^{-8}	10^{-7}	10^{-6}	10^{-5}	10^{-4}
10^{-8}	Précision non atteinte	Précision non atteinte	I=41 0,291 666 62 0,083 333 32 0,083 333 32 0,041 666 62	I=35 0,291 666 0 0,083 333 0 0,083 333 0 0,041 666 0	I=29 0,291 664 0,083 328 0,083 328 0,041 664
10^{-7}	I=44 0,291 66 65 0,083 333 31 0,083 333 31 0,041 666 65	I=43 0,291 666 62 0,083 333 31 0,083 333 31 0,041 666 62 K>48	I=41 -	I=35 -	I=29 -
10^{-6}	I=37 0,291 666 5 0,083 333 0 0,083 333 0 0,041 666 5	I=37 -	I=37 -	I=33 0,291 666 0 0,083 332 0 0,083 332 0 0,041 666 0	I=29 -
10^{-5}	I=31 0,291 664 0,083 332 0,083 332 0,041 664	I=31 -	I=31 -	I=31 -	I=27 0,291 656 0,083 328 0,083 328 0,041 656

III Méthode associée à la décomposition d'une norme [15]

III.1 Principe

Soit une norme p décomposée comme au chapitre I (§ I.5.4.)

$$p(x) = z^t(x).x$$

On passera de x_m à x_{m+1} par la relation de récurrence

$$x_{m+1} = x_m = \frac{p(r_m)}{||A_\bullet^t\ z(r_m)||^2}\ A_\bullet^t\ z(r_m)$$

La norme euclidienne est notée ici || ||.

N.Gastinel a démontré que ce processus était toujours convergent quelque soit le vecteur x_1 de départ.

Nous ferons ici une étude pratique dans le cas où la norme p est la norme de Holder d'ordre 1.

Nous avons :

$$x_{m+1} = x_m - \frac{H_1(r_m)}{||A^t\ z(r_m)||^2}\ A_\bullet^t\ z(r_m)$$

$$r_{m+1} = r_m - \frac{H_1\ (r_m)}{||A^t\ z(r_m)||^2}\ AA^t\ z(r_m)$$

III - 2 Etude pratique

1. On prendra ici les normes S_{11} et H_1
Les lieux d'incertitude seront :

$$S_{11}(A) = \max_j \sum_i |a_{ij}| \leq \alpha$$

$$H_1(b) = \sum_i |b_i| \leq \beta$$

2. Procédure ALGOL

```
procedure RESOLTEST3(A)SECOND MEMBRE:(B) ORDRE:(N) VECTEUR
          SOLUTION:(X) INCERTITUDE SUR A:(ALFA) INCERTITUDE
          SUR B:(BETA) NORME X:(Q) NORME R:(P) ;
entier N; reel ALFA, BETA, P, Q; reel tableau A, B, X ;
commentaire Résolution par une méthode itérative (Décomposition
d'unenorme -Gastinel-) d'un système linéaire dont les données
(A,B) sont considérées comme incertaines (Norme de R :P=H1,
Norme de X : Q=H1) ;
début procedure GASTINEL(A) SECOND MEMBRE:(B) ORDRE:(N)
                VECTEUR ITERE:(X) NORME R:(P) NORME X:(Q);
entier N; reel P, Q ; reel tableau A, B, X ;
commentaire Passage d'un vecteur itéré à l'itéré suivant par
la méthode de Gastinel -(décomposition de la norme H1)- ;
```

```
      debut reel S, L, C; reel tableau R, Z, Y[1:N];
            P := 0.0;   L := 0.0;
            pour I := 1 pas 1 jusqua N faire
            debut  S := 0.0;
                   pour J := 1 pas 1 jusqua N faire
                   S := S + A[I,J]x X[J];
                   R[I]:= S - B[I] ;
                   Z[I]:= SIGNE (R[I]);
                   P := P + ABS (R[I])
            fin
            pour I := 1 pas 1  jusqua N faire
            debut  S := 0.0;
                   pour J := 1 pas 1 jusqua N faire
                   S := S + A[J,I]x Z[J];
                   Y[I] := S ;
                   L := L + S↑2
            fin
            C := P/L ;
            Q := 0.0;
            pour I := 1 pas 1 jusqua N faire
            debut  X[I]:= X[I]- C x Y[I];
                   Q := Q + ABS (X[I])
            fin
      fin   GASTINEL
      reel E;
BOUCLE :    GASTINEL (A, B, N, X, P, Q);
            E := ALFA x Q + BETA;
            si P>E alors allera BOUCLE
 fin RESOLTEST 3 ;
```

3. Résultats numériques

Exemple 1

$$A_o = \begin{vmatrix} 4,33 & -1,12 & -1,08 & 1,14 \\ -1,12 & 4,33 & 0,24 & -1,22 \\ -1,08 & 0,24 & 7,21 & -3,22 \\ -1,14 & -1,22 & -3,22 & 5,43 \end{vmatrix} \qquad b_o = \begin{vmatrix} 3,52 \\ 1,57 \\ 0,54 \\ -1,09 \end{vmatrix}$$

β \ α	10^{-7}	10^{-6}	10^{-5}
10^{-7}	I=79 1,046 249 03 0,562 782 88 0,111 002 70 -0,228 121 57 K=88	I=72 1,046 249 15 0,562 783 00 0,111 002 72 -0,228 121 34	I= 63 1,046 249 50 0,562 783 24 0,111 002 93 -0,228 119 23
10^{-6}	I = 74 1,046 249 03 0,562 782 94 0,111 002 71 -0,228 121 50	I=72 - K=75	I=63 -
10^{-5}	I=65 1,046 249 62 0,562 783 18 0,111 003 12 -0,228 119 90	I=65 -	I=63 - K=67

Exemple 2

$$A_o = \begin{vmatrix} 0{,}734\ 841 & 0{,}262\ 827 & 0{,}013\ 905 \\ 0{,}262\ 827 & 1{,}052\ 633 & 0{,}337\ 503 \\ -0{,}013\ 905 & -0{,}337\ 503 & 1{,}331\ 937 \end{vmatrix} \qquad b_o = \begin{vmatrix} 1{,}218\ 78 \\ 1{,}355\ 84 \\ 3{,}306\ 9 \end{vmatrix}$$

β \ α	10^{-7}	10^{-6}	10^{-5}	10^{-4}
10^{-7}	I=55 1,000 000 23 1,999 999 76 2,999 999 76 K=61	I=47 1,000 002 2 2,000 000 7 2,999 999 7	I=38 1,000 029 1,999 989 2,999 984	I=28 1,000 24 2,000 06 2,999 97
10^{-6}	I=51 1,000 000 95 2,000 000 00 2,999 999 76	I=45 1,000 005 1 1,999 997 1 2,999 996 6 K=55	I=38 -	I=28 -
10^{-5}	I=43 1,000 007 3 1,999 997 8 2,999 996 6	I=42 1,000 003 4 1,999 995 2 3,000 000 4	I=36 1,000 043 1,999 994 2,999 984 K=45	I=28 -
10^{-4}	I=34 1,000 065 2,000 002 2,999 983	I=34 -	I=32 1,000 100 2,000 015 2,999 982	I=26 1,000 55 1,999 70 2,999 64 K=34

Exemple 3

$$A_o = \begin{vmatrix} 3 & 2 & 1 \\ 5 & -6 & 1 \\ -4 & 2 & 1 \end{vmatrix} \qquad b_o = \begin{vmatrix} 10 \\ -4 \\ 3 \end{vmatrix}$$

β \ α	10^{-7}	10^{-6}	10^{-5}	10^{-4}
10^{-7}	Précision non atteinte	I=54 0,999 999 76 1,999 999 52 2,999 999 76	I=44 1,000 000 11 1,999 994 5 2,999 997 6	I=33 1,000 004 7 1,999 967 8 2,999 965 6
10^{-6}	I=59 1,000 000 00 1,999 999 88 2,999 999 76	I=54 - K=59	I=44 -	I=33 -
10^{-5}	I=52 1,000 000 35 1,999 999 88 2,999 999 04	I=50 1,000 000 00 1,999 998 2 2,999 999 2	I= 42 1,000 001 6 1,999 992 2,999 998 K=51	I=33 -
10^{-4}	I=40 1,000 005 1,999 994 2,999 994	I=40 -	I=38 1,000 000 9 1,999 983 7 2,999 993 5	I=33 - K=36

Exemple 4

Pour A_o et b_o se référer à la page III.21

β \ α	10^{-7}	10^{-6}	10^{-5}	10^{-4}
10^{-7}	Précision non atteinte	I=153 0,299 107 10 0,098 214 26 0,031 249 98 0,098 214 26 0,062 499 94 0,026 785 70 0,031 249 98 0,026 785 70 0,013 392 81	I=129 0,299 106 65 0,098 213 75 0,031 249 64 0,098 213 79 0,062 499 12 0,026 785 01 0,031 249 64 0,026 785 01 0,013 392 56	I=105 0,299 103 64 0,098 214 29 0,031 248 25 0,098 214 29 0,062 491 29 0,026 780 81 0,031 248 25 0,026 780 81 0,013 390 68
10^{-6}	I=146 0,299 107 01 0,098 214 20 0,031 249 96 0,098 214 20 0,062 499 83 0,026 785 64 0,026 249 96 0,026 785 64 0,013 392 83	I=142 0,299 107 04 0,098 214 15 0,031 249 95 0,098 214 15 0,062 499 80 0,026 785 60 0,026 249 95 0,026 785 60 0,013 392 86 K=146	I=123 0,299 106 68 0,098 213 70 0,031 249 30 0,098 213 70 0,062 498 45 0,026 784 96 0,026 249 30 0,026 784 96 0,013 392 26	I=105 -
10^{-5}	I=120 0,299 106 35 0,098 213 32 0,031 249 72 0,098 213 32 0,062 498 21 0,026 784 83 0,026 249 72 0,026 784 53 0,013 392 86	I=120 -	I=117 0,299 106 68 0,098 213 22 0,031 249 63 0,098 213 22 0,062 497 62 0,026 785 27 0,026 249 03 0,026 785 27 0,013 392 24 K=123	I=105 -

Exemple 5

$$A_o = \begin{vmatrix} -13 & 2 & -1 & 3 \\ -6 & 19 & 4 & 5 \\ 1 & -5 & -37 & 7 \\ 5 & 9 & 8 & 23 \end{vmatrix} \qquad b_o = \begin{vmatrix} -61 \\ -217 \\ 247 \\ 590 \end{vmatrix}$$

α / β	10^{-7}	10^{-6}	10^{-5}	10^{-4}
10^{-7}	Précision non atteinte	I=120 7,000 003 38 -14,000 000 00 - 0,999 999 58 19,000 000 00	I=107 7,000 013 3 -13,999 997 1 - 1,000 003 9 19,000 000 0	I=87 7,000 123 -13,999 952 - 0,999 985 19,000 017
10^{-6}	Précision non atteinte	I=120 - K>150	I=107 -	I=87 -
10^{-5}	Précision non atteinte	I=120 -	I=107 - K=140	I=87 -
10^{-4}	II=115 7,000 004 2 -13,999 999 0' - 1,000 000 95 18,999 998 09	I=114 7,000 004 7 -19,999 999 04 - 0,999 999 40 18,999 998 09	I=105 7,000 016 6 -13,999 998 0 - 1,000 004 0 19,000 000 00	I=87 - K=115

B I B L I O G R A P H I E

----- oOo -----

[1] J. von NEUMANN - *Numerical inverting of matrices of high order.* Bul. Amer. Math.Soc., vol.53, n°11, pp.1021-1099 (1947)

[2] A. M. TURING - *Rouding-off errors in matrix processes.* Quart. J. mech. appl. math. I (1948), pp. 287-308

[3] N. GASTINEL - *Matrices du Second Degré et normes générales en Analyse Numérique Linéaire.* Publications scientifiques et techniques du Ministère de l'Air (1962), chap. III et IV

[4] N. BOURBAKI - *Topologie générale.* chap. IX

[5] A.S. HOUSEHOLDER - *The Theory of Matrices in Numerical Analysis.* chap. II, § 2.1 - § 2.2

[6] N. GASTINEL - *Matrices du Second Degré et normes générales en Analyse Numérique Linéaire.* Publications scientifiques et techniques du Ministère de l'Air (1962), chap. VII. § 3

[7] A.M. OSTROWSKI - *Uber Normen von Matrizen.* Math. Zeitsch., Bd 69, Sp 2 (1955).

[8] N. GASTINEL - *Matrices du Second Degré et normes générales en Analyse Numérique Linéaire.* Publications scientifiques et techniques du Ministère de l'Air (1962), chap. II. § 3

[9] F. ROBERT - *Normes vectorielles de vecteurs et de matrices.* Revue française du traitement de l'Information Chiffres - 4 (1964)

[10] J.L. RIGAL - FRANCK - *Sur la distance minimale d'une matrice régulière donnée au lieu des matrices singulières.* Actes du II^{e}. Congrès l'A.F.C.A.L.T.I. (octobre 1961 p. 55)

[11] W. OETTLI - W. PRAGER - *Compatibility of Approximate Solution of Linear Equations with Given Error Bounds for Coefficients ans Right-Hand Sides.* Numerische Mathematik 6, 405-409 (1964)

W. OETTLI - W. PRAGER - J. H. WILKINSON - *Admissible Solutions of Linear systems with not sharply defined coefficients.* J. SIAM Numer. Anal. ser. B. Vol. 2, No. 2 (1965)

[12] NGUYEN HUU VINH - *Semi-norme duale généralisée et approximation d'un vecteur.* C.R. Acad. Sc. Paris t.262 pp. 1456 - 59 (1966)

[13] N. GASTINEL - *Analyse Numérique Linéaire.* Edit. Hermann pp. 135-140

[14] GAUSS - *Gottingen.* 1823. Werke IV (1873)

SEIDEL - *Denkschrifte r der Munchener Akademie* (1862); *Abhandlungen Bayeriche Akad. Wissenchaften,* Band 11 (1874), pp. 81-108.

[15] N. GASTINEL - *Matrices du Second Degré et normes générales en Analyse Numérique Linéaire.* Publications scientifiques et techniques du Ministère de l'Air (1962), chap. VII § 4 - 6

[16] J. GACHES - "*Compatibilité d'une solution* $\tilde{x}$ *avec un système linéaire donné*" . C.R. Acad. Sc. Paris, t.262 pp. 348-349 (1966)

JURY : MM. CHATELET Président

GASTINEL

RIGAL Examinateurs

PARIZET

VU et APPROUVE
Besançon , le

Le Doyen de la Faculté des Sciences

VU, et permis d'imprimer,
Besançon, le

Le Recteur de l'Académie de Besançon